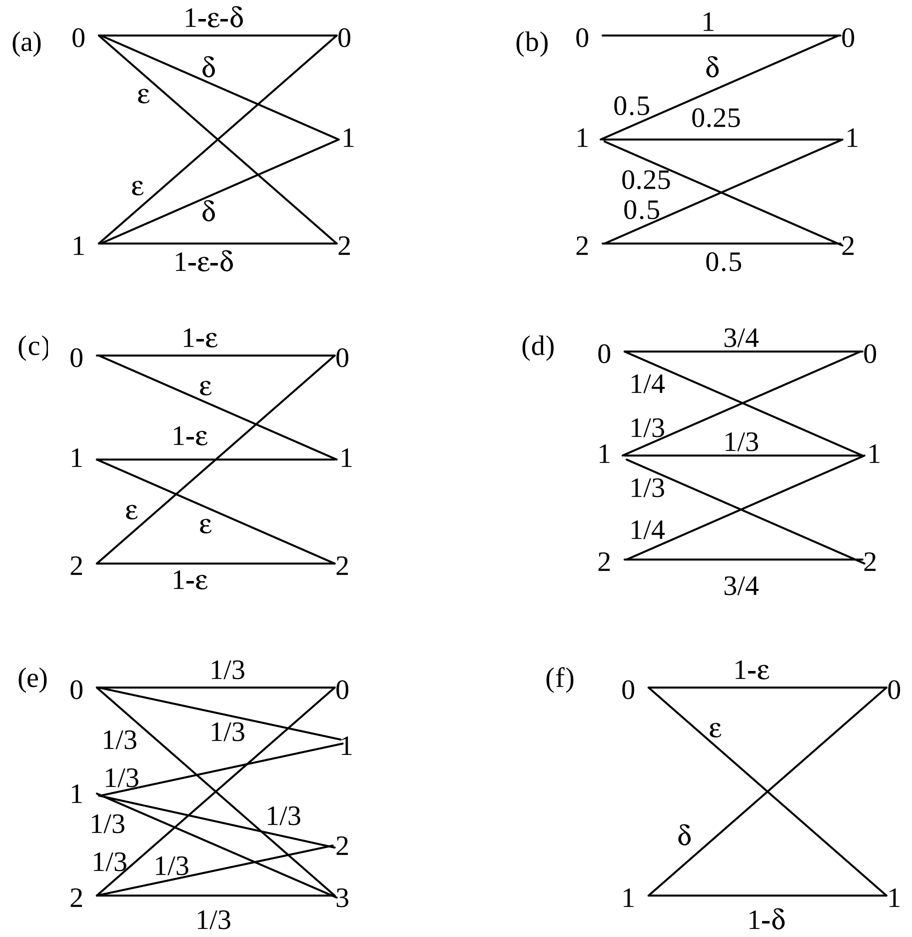
《信息论与编码》第四章习题解答

4.1 **计算如下所示离散无记忆信道的容量：**



习题 4.1 **图**

[解]

（a）**信道概率转移矩阵为**

$$P=\begin{pmatrix}1-e-d & d & e\\ e & d & 1-e-d\end{pmatrix},$$

信道是准对称信道，因此在输入为等概分布时达到信道容量，即 $P(X=0)=P(X=1)=0.5$ **时达到信道容量。这时**

$$P(Y=0)=0.5-0.5d$$

$$P(Y=1)=d$$

$$P(Y=2)=0.5-0.5d$$

相应的信道容量为

$$C=I(X=0;Y)=I(X=1;Y)$$

$$=\sum_{j=0}^{2}p(j\,|\,0)\log\frac{p(j\,|\,0)}{p(j)}$$

$$= (1-e-d)\log\frac{1-e-d}{0.5-0.5d} + d\log\frac{d}{d} + e\log\frac{e}{0.5-0.5d}$$
$$= (1-e-d)\log(1-e-d) + e\log e - (1-d)\log(0.5-0.5d)$$

(b) **信道概率转移矩阵为**

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

当 $P(X=0)=P(X=2)=0.5$ **，** $P(X)=0$ **时，**

$P(Y=0)=0.5$ **，** $P(Y=1)=0.25$ **，** $P(Y=2)=0.25$

$$I(X=0;Y)=\sum_{j=0}^{2} p(j|0)\log\frac{p(j|0)}{p(j)}=1 \text{ bit}$$

$$I(X=2;Y)=\sum_{j=0}^{2} p(j|2)\log\frac{p(j|2)}{p(j)}$$
$$=0.5\log\frac{0.5}{0.25}+0.5\log\frac{0.5}{0.25}=1 \text{ bit}$$

$I(X=1;Y)=0\le 1$ **；**

所以满足定理 4.2.2 **条件，由达到信道容量充要条件可知，信道容量**

C=1 bit/**次**

(c) **信道转移概率矩阵为**

$$P=\begin{pmatrix} 1-e & e & 0 \\ 0 & 1-e & e \\ e & 0 & 1-e \end{pmatrix},$$

信道是对称信道，当输入为均匀分布时，即

$$P(X=0)=P(X=1)=P(X=2)=\frac{1}{3}$$

时，达到信道容量。

$$C=\log 3+e\log e+(1-e)\log(1-e)$$
$$=\log 3-H(e,1-e)$$

(d) **信道转移概率矩阵为**

$$P=\begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

当 $p(X=0)=p(X=2)=0.5$ **，** $p(X=1)=0$ **时，**

$P(Y=0)=P(Y=2)=\frac{3}{8}$ **，** $P(Y=1)=\frac{2}{8}$

$$I(X=0;Y)=\sum_{j=0}^{2} p(j|0)\log\frac{p(j|0)}{p(j)}$$

$$= \frac{3}{4}\log 2 + \frac{1}{4}\log 1$$

$$= \frac{3}{4} \text{ bit}$$

$$I(X=2;Y) = \frac{3}{4} \text{ bit}$$

$$I(X=1;Y) = \sum_{j=0}^{2} p(j|1)\log\frac{p(j|1)}{p(j)}$$

$$= \frac{1}{3}\log\frac{1/3}{3/8} + \frac{1}{3}\log\frac{1/3}{2/8} + \frac{1}{3}\log\frac{1/3}{3/8}$$

$$= \frac{2}{3}\log\frac{8}{9} + \frac{1}{3}\log\frac{8}{6}$$

$$\le \frac{3}{4} \text{ bit}$$

所以满足定理 4.2.2 **所规定的达到信道容量的充要条件，信道容量为**

$$C = \frac{3}{4} \text{ bit/次}$$

（e）

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

信道是准对称信道，当输入分布为均匀分布时达到信道容量，即

$p(X=0) = p(X=1) = p(X=2) = \frac{1}{3}$ **时达到信道容量。信道容量为**

$$C = \sum_{i=0}^{2}\sum_{j=0}^{3} p(i)p(j|i)\log\frac{p(j|i)}{p(j)}$$

其中 $p(Y=0) = p(Y=1) = p(Y=2) = \frac{2}{9}$

$$p(Y=3) = \frac{1}{3}$$

所以

$$C = 6\cdot\frac{1}{9}\log\frac{1/3}{2/9} + 3\cdot\frac{1}{9}\log\frac{1/3}{1/3}$$

$$= \frac{2}{3}\log\frac{3}{2} \text{ bit/次}$$

（f）**信道转移概率矩阵**

$$P = \begin{pmatrix} 1-e & e \\ d & 1-d \end{pmatrix}$$

利用方程求逆方法计算信道容量。设

$p(X=0) = q$，$p(X=1) = 1-q$，$0 < q < 1$

则 $w_0 = p(Y=0) = q(1-e) + (1-q)d$

$w_1 = p(Y=1) = q \cdot e + (1-q)\cdot(1-d)$

记 $b_0 = C + \log w_0$

$b_1 = C + \log w_1$

于是得到

$$(1-e)b_0 + eb_1 = (1-e)\log(1-e) + e\log e = -H(e)$$

$$db_0 + (1-d)b_0 = d\log d + (1-d)\log(1-d) = -H(d)$$

解上面方程得

$$b_0 = \frac{eH(d)-(1-d)H(e)}{1-d-e}$$

$$b_1 = \frac{dH(e)-(1-e)H(d)}{1-d-e}$$

所以

$$C = \log\sum_j 2^{b_j}$$

$$= \log\left[2^{\frac{eH(d)-(1-d)H(e)}{1-d-e}} + 2^{\frac{dH(e)-(1-e)H(d)}{1-d-e}}\right]$$

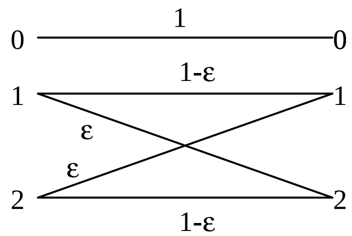
由 $w_0 = 2^{b_0 - C}$，$w_1 = 2^{b_1 - e}$

解出

$$q = \frac{(1-d)w_0 - dw_1}{1-e-d}$$

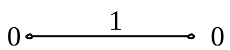
$$1-q = \frac{(1-e)w_1 - ew_0}{1-e-d}$$

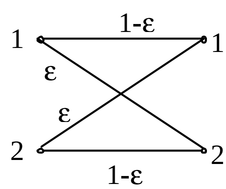
4.2 **计算如下信道的容量，**



习题 4.2 **图**

[解] 把该信道看成是退化信道C_1**和二元对称信道**C_2 **的和信道，**





退化信道C_1　　　　**二元对称信道**C_2

退化信道容量为 $C_1=0$，二元对称信道容量为 $C_2=1-H(e)$，

所以和信道的容量为

$$C=\log\left[1+2^{1-H(e)}\right]$$

达到信道容量的输入分布为

$$\begin{aligned}p(X=0)&=2^{C_1-C}\\&=\frac{1}{1+2^{1-H(e)}}\\p(X=1)&=p(X=2)\\&=0.5\cdot 2^{C_2-C}\\&=\frac{2^{-H(e)}}{1+2^{1-H(e)}}\end{aligned}$$

4.3 考虑离散无记忆信道 $Y=X+Z$（mod 11），

其中 $$Z=\begin{pmatrix}1, & 2, & 3\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{pmatrix} \qquad X\in\{0,1,\cdots,10\}$$

假定 Z 和 X 独立，求：

（1）信道容量，

（2）达到信道容量的输入分布 $\{P^*(x)\}$。

[解] 信道为对称信道，所以当输入为均匀分布时，即

$$p(X=i)=\frac{1}{11}\text{，}\ i=0,1,\cdots,10$$

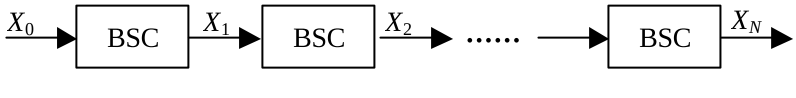
时达到信道容量。信道容量为

$$\begin{aligned}C&=\log 11-H(\frac{1}{3},\frac{1}{3},\frac{1}{3})\\&=\log\frac{11}{3}\ \text{bit/次}\end{aligned}$$

4.4 用差错率为 e 的二元对称信道 BSC 构成如下复合信道

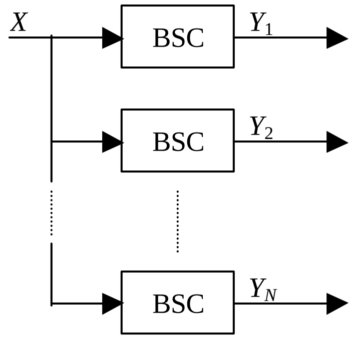
（1）长度为 N 的级联信道

求该级联信道的容量 C_N，并证明 $\lim\limits_{N\to\infty} C_N=0$

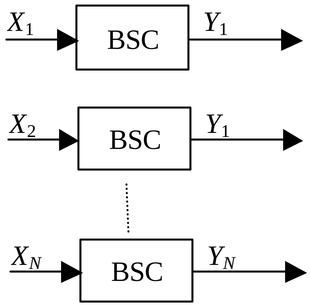


习题 4.4（1）图　级联信道

（2）并联输入信道，把输入 X 并联接到各信道，输出是矢量，当 $N\to\infty$ 时并联输入信道容量趋于 1。



习题 4.4（2）图　并联输入信道



习题 4.4（3）图　积信道

（3）N **个相同** BSC **的积信道，求这时积信道容量** C_N **，且证明** $\lim\limits_{N\to\infty} C_N = \infty$

[证明]

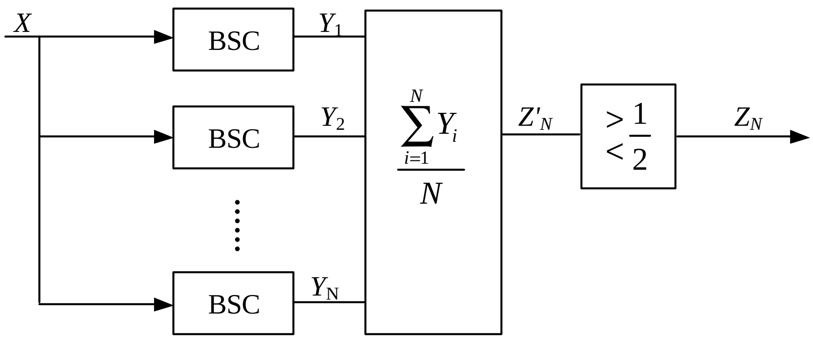
（1）**见例** 4.3.2

（2）**首先因为**

$$\begin{aligned} I(X;Y_1,Y_2,\cdots,Y_N) &= H(X) - H(X \mid Y_1,Y_2\cdots Y_N) \\ &\le H(X) \\ &\le 1 \ \text{bit} \end{aligned}$$

所以对任何 N**，并联输入信道容量** $C_N \le 1$ bit**，下面证明当** $N \to \infty$ **时，** $C_N \ge 1$ bit。

我们对输出 $Y_1,Y_2,\cdots Y_N$ **作进一步处理，如下图所示**



其中　　　$Z'_N = \sum_{i=1}^{N} Y_i / N$

$$Z_N=\begin{cases}0 & Z'_N\le\frac{1}{2}\\ 1 & Z'_N>\frac{1}{2}\end{cases}$$

由于当 X=0 时，Y_i 是独立，同分布二元随机变量

$$p(Y_i=1\mid X=0)=p\text{，}\ p(Y_i=0\mid X=0)=1-p$$

$$E[Z'_N\mid X=0]=p\text{，}\ s^2_{Z'_N|X=0}=\frac{p(1-p)}{N}$$

利用切比雪夫不等式

$$\begin{aligned}P[Z_N=1\mid X=0]&=P\left\{Z'_N>\frac{1}{2}\mid X=0\right\}\\&=P\left\{Z'_N-p>\frac{1}{2}-p\mid X=0\right\}\\&\le P\left\{|Z'_N-p|>\frac{1}{2}-p\mid X=0\right\}\\&\le\frac{s^2_{Z'_N|X=0}}{\left(\frac{1}{2}-p\right)^2}\\&=\frac{p(1-p)}{N(\frac{1}{2}-p)^2}\end{aligned}$$

当 $p<\frac{1}{2}$**，以及** N **充分大时**

$$P[Z_N=1\mid X=0]\to 0$$

类似的当 N **充分大时**

$$P[Z_N=0\mid X=0]\to 1$$
$$P[Z_N=0\mid X=1]\to 0$$
$$P[Z_N=1\mid X=1]\to 1$$

于是当 $N\to\infty$ **时，**X **和** Z_N **之间的等效信是趋于无噪信道，即**

$$\lim_{N\to\infty}\max_{\{p(x)\}}I(X;Z_N)=1\ \text{bit}$$

由于数据处理定理

$$I(X;Y_1\cdots Y_N)\ge I(X;Z_N)$$

所以 $\lim_{N\to\infty}C_N\ge 1\ \text{bit}$**，**

从而 $\lim_{N\to\infty}C_N=1\ \text{bit}$

（3）积信道容量

$$C_N=N\cdot C$$

其中 $C=1-H(p)$

所以 $\lim_{N\to\infty}C_N=\infty$

4.5 **离散无记忆信道由下述矩阵描述**

	y_1	y_2	y_3
x_1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$
x_2	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$
x_3	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

如 $p(x_1)=\frac{1}{2}$**，** $p(x_2)=p(x_3)=\frac{1}{4}$**，求理想观察员判别方式，即最大后验概率判决方式，以及相应的错误概率。**

[解] 理想观察员判别方式就是最大后验概率判别方式，即如果接收到 $Y=j$**，则理想观察员按如下作出判别**

$$\hat{x}(Y=j)=\arg\max_{x\in X}\ p(x\,|\,Y=j)$$
$$=\arg\max_{x\in X}\frac{p(x)\cdot p(Y=j\,|\,x)}{p(Y=j)}$$

于是对于每个 $Y=j$**，** $j=1,2,3$ **可以计算出** $\hat{x}(Y=j)$ **为：**

y	y_1	y_2	y_3
$\hat{x}(y)$	x_1	x_1	x_3

错误概率 $P_e=P\{X\neq\hat{X}(Y)\}$

$$=\sum_x p(x)\cdot p\{\hat{x}\neq x\,|\,x\}$$
$$=p(x_1)\cdot p(y_3\,|\,x_1)+p(x_2)\cdot 1+p(x_3)\cdot[p(y_1\,|\,x_3)+p(y_2\,|\,x_3)]$$
$$=\frac{1}{2}\cdot\frac{1}{6}+\frac{1}{4}+\frac{1}{4}\cdot\frac{1}{2}$$
$$=\frac{11}{24}$$

4.6 **信道输入，输出字符集相同** $X=Y=\{0,1,2,3,4,\}$**，信道转移概率为**

$$p(y\,|\,x)=\begin{cases}\frac{1}{2} & \text{如}\ y=x\pm 1 \bmod 5\\ 0 & \text{其它}\end{cases}$$

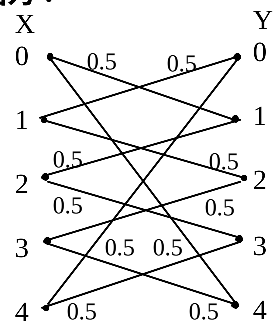
（1）求信道容量；

（2）信道的“0—错误”容量是指每次信道以 0 错误概率传输的最大比特数。显然对于本信道来说“0—错误”容量至少为 1 比特，（因为若以 $\frac{1}{2}$ 概率传输符号“0”和“1”可达“0—错误”传输）。请找出“0—错误”传输速率大于 1 比特的分组码？你能否估计该信道“0—错误”容量？

[解]（1）转移概率矩阵为

$$P=\begin{pmatrix} 0 & 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 & 0 \end{pmatrix}$$

相应的转移概率图为：



由转移概率矩阵看出信道是对称信道，所以，在输入为均匀分布时达到信道容量

$$C=\log 5+0.5\log 0.5+0.5\log 0.5$$

$$=\log\frac{5}{2}\quad \text{bit/次}$$

（2）如果有二个 $i\neq j$，使得存在 k 使

$$p(Y=k\mid X=i)\neq 0\text{，以及 }P(Y=k\mid X=j)\neq 0$$

则称 i 和 j 相邻。如果选择彼此不相邻的符号传输，是不会发生错误的，这就是“0—错误”传输。显然 $X=0$ 和 $X=1$ 是二个不相邻的符号，于是如果采用 0.5 概率传 $X=0$，0.5 概率传 X=1 则可以达到以“0—错误”传送一个比特的码率。

可以构成长度为 2 的码字，这时有 25 种不同输入 (x_1,x_2)，$x_1,x_2\in\{0,1,2,3,4\}$，和 25 种不同输出 (y_1,y_2)，$y_1,y_2\in\{0,1,2,3,4\}$。对于长度为 2 的扩展符号来说，可以构成 25×25 转移概率矩阵和相应的转移概率图，发现（0,0），（2,4），（4,3），（1,4），（3,1）是 5 个不相邻的扩展符号，等概率使用这 5 个扩展符号传输信息可以得到“0—错误”传输，而且平均每次使用信道（每符号）传输信息速率为

$$R=\frac{1}{2}\log 5\quad \text{bit/次}$$

于是“0—错误”容量

$$C_0\geq\frac{1}{2}\log 5\quad \text{bit/次}$$

显然“0—错误”容量一定小于一般容量，所以下式应该成立

$$C_0\leq\log\frac{5}{2}\quad \text{bit/次}$$

4.7 给定一个信道 $\{p(y_j\mid x_i)\}_{MXL}$，输入、输出字符表分别为 $\{x_1,x_2,\cdots x_M\}$，$\{y_1,y_2,\cdots,y_L\}$。一个随机判决方式如下，对一个信道输出 $y_j(j=1,2,\cdots,L)$，译码器以概率 q_{ji} 选定符号 $x_i(i=1,2,\cdots,M)$，试证明对一个给定的输入分布，没有一种随机判决方式具有比理想观察员方式有更小的错误概率。

[证明]如果接收到$Y=y_j$，则由随机判决方式正确传输概率为

$$P_C(y_j)=\sum_{i=1}^{M}q_{ji}P(X=x_i\mid y_j)$$

如采用理想观察员方式，即

$$\hat{x}(y_j)=\arg\max_i P(X=x_i\mid y_j)$$

或者说

$$P\{X=\hat{x}(y_j)\mid y_j\}=\max_i P\{X=x_i\mid y_i\}$$

则正确传输概率为

$$\begin{aligned}\hat{P}_C(y_j)&=P\{X=\hat{x}(y_j)\mid y_j\}\\&=\sum_i q_{ji}\cdot P\{X=\hat{x}(y_j)\mid y_j\}\\&\ge\sum_i q_{ji}\cdot p\{X=x_i\mid y_i\}\\&=P_C(y_j)\end{aligned}$$

所以理想观察员方式具有比任何随机判决方式更大的正确判决概率。

4.9 **计算下述信道容量**

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

[解] 本信道是二个相同的二元对称信道的和信道，每个分量信道的容量为

$$C_1=C_2=1-H(p)$$

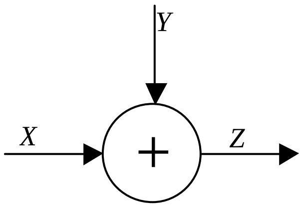
在输入为均匀分布时达到容量。于是和信道容量为

$$C=\log[2^{C_1}+2^{C_2}]=2-H(p)\quad \text{bit/次}$$

相应输入均匀分布为：

$$p(x_0)=p(x_1)=p(x_2)=p(x_3)=\frac{1}{4}$$

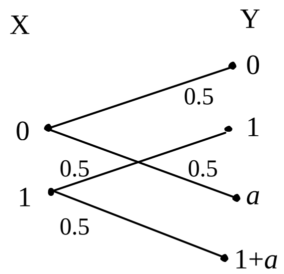
4.10 **离散无记忆加性噪声信道如图所示。其输入随机变量** X **和噪声** Y **统计独立。**X **的取值为**{0，1}，Y **取值为**{0，a}（$a\ge1$），**又** $p(y=0)=p(y=a)=0.5$，**信道输出** $Z=X+Y$，**求此信道容量，以及达到信道容量的最佳输入分布。注意信道容量依赖于** a **的取值。**



习题 4.10 图

［解］ 由题意 $Z = X + Y$，当 $a > 1$ 时相应的转移概率图与转移概率矩阵为

$$\begin{array}{ccccc} & 0 & 1 & a & 1+a \\ & \downarrow & \downarrow & \downarrow & \downarrow \\ P = & \multicolumn{4}{c}{\begin{pmatrix} 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \end{pmatrix}} \end{array}$$

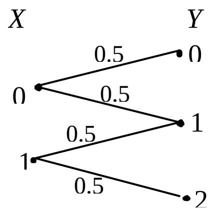


这时为无损信道，容量为

C=1 bit/次

当 a=1 时，相应转移概率图和转移概率矩阵为

$$P = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$



这时为二元除删信道，容量为

$$C = 1 - 0.5 = 0.5 \quad \text{bit/次}$$

4.14 给定系统带宽为 W，噪声功率谱密度为 N_0，试证明传送一比特信息所需最小能量为 $0.693\,N_0$（瓦）。如果要求 $\frac{S}{WN_0} > 4000$，其中 S 为信号功率，试证明所需的信号能量至少为此最小值的 482 倍。

［证明］由 4.6.1 节可知在频带效率 h 下，每传 1bit 所需能量为

$$E_b(h) \geq \frac{2^h - 1}{h} \cdot N_0$$

其中不等式右边是 $h = \frac{R_b}{W} = \frac{1}{WT_b}$ 的增函数。对固定 W，当 $T_b \to \infty$ 时，$h \to 0$，这时达到最小值

$$E_{\min} = \lim_{h \to 0} E_b(h) = 0.693 N_0$$

设用 T_b 时间传送 1bit 信息，由 Shannon 公式

$$C_{T_b} = T_b \cdot W \log \cdot \left(1 + \frac{S}{N_0 W}\right) = 1 \quad \text{bit}$$

得到
$$T_b W = 1 / \log\left(1 + \frac{S}{N_0 W}\right)$$

当 $\frac{S}{N_0 W} > 4000$ 时，由于 $S = E_b / T_b$，所以

$$E_b > 4000 N_0 \cdot W T_b = \frac{4000 N_0}{\log\left(1+\frac{S}{N_0 W}\right)}$$

因此
$$\frac{E_b}{E_{\min}} \geq \frac{4000 N_0}{0.693 \cdot N_0 \cdot \log(4001)} = 482$$

4.11 信道输入，输出在实数轴上取值，即 $\mathrm{X}=\Upsilon=R, Y=X+Z$，Z 为 $[-0.5,0.5]$ 上均匀分布随机变量，X 的峰值限制为 $|X|\leq\beta$， β 为正整数，求证信道容量为：

$$C(\beta)=\log(2\beta+1)$$

[证明] Y 均匀取值为 $[-\beta-0.5, 0.5+\beta]$ 时 $H(Y)$ 有最大值。

$$\begin{aligned}I(X,Y)&=H(Y)-H(Y|X)\\&=H(Y)-(Z)\\&\leq\int_{-\beta-0.5}^{0.5+\beta}\frac{1}{2\beta+1}kog(2\beta+1)dx-\int_{-0.5}^{0.5}\log(1)dx\\&=\log(2\beta+1)\end{aligned}$$

即 $C(\beta)=\log(2\beta+1)$

4.12 有一个叠加噪声的信道，输入 X 是离散随机变量，取值为 $\{1,-1\}$，噪声 N 的概率密度为 $P_N(n)=\begin{cases}\frac{1}{4} & |n|\leq 2\\ 0 & |n|>2\end{cases}$，信道输出 Y＝X＋Z 是连续随机变量，求：

（1） 该半连续信道容量 $C=\max\limits_{\{p(x)\}}I(X;Y)$

（2） 若在信道输出上接一个检测器作为这个信道的一部分，检测器输出变量为：

$$Z=\begin{cases}1 & Y>1\\ 0 & -1\leq Y\leq 1\\ -1 & Y<-1\end{cases}$$

这样构成离散信道，求它的容量。

[解] （1） $C=\max(H(Y)-H(Y|X))$

显然信道对称，此时输入等概时，即 $p(x=-1)=p(x=1)=\frac{1}{2}$ 时，达到信道容量

此时的输出为：

$$p(Y)=\begin{cases}\frac{1}{4} & 1\leq Y\leq 3\\ \frac{1}{2} & -1\leq Y\leq 1\\ \frac{1}{4} & -3\leq Y\leq -1\end{cases}$$

概率密度为：

$$p(y)=\begin{cases}\frac{1}{8} & 1\le y\le 3\\ \frac{1}{4} & -1\le y\le 1\\ \frac{1}{8} & -3\le y\le -1\end{cases}$$

则：$C=H(Y)-H(Y|X)$

$$\le\int_1^3\frac{1}{8}\log(8)dx+\int_{-1}^1\frac{1}{4}\log(4)dx+\int_{-3}^{-1}\frac{1}{8}\log(8)dx-\int_{-2}^{2}\frac{1}{4}\log(4)dx$$

＝0.5bit

（2） Z 的概率为：

$$p(Z)=\begin{cases}\frac{1}{4} & Z=1\\ \frac{1}{2} & Z=0\\ \frac{1}{4} & Z=-1\end{cases}$$

则：$C=H(Z)-H(Z|X)$

$$=H(Z)-H(Z|X,Y)-H(Y|X)$$

$$=\frac{1}{4}\log 4+\frac{1}{2}\log 2+\frac{1}{4}\log 4-0-1$$

＝0.5bit

4.13 设一个连续信道，传送信息 $X\in(-\pi,\pi)$，信道受到加性高斯白噪声的干扰，而 Z 的概率密度为：

$$p(z)=\begin{cases}\frac{1}{2a} & |z|\le a, a>0\\ 0 & |z|>a\end{cases}$$

输出 $Y=(X+Z)\bmod 2\pi$，试求：

（1） $a>\pi$ 时的信道容量。

（2） $a\le\pi$ 时的信道容量。

[解] $C=\max(I(X,Y))=\max(H(Y)-H(Y/X))$

令 $Y'=X+Z$ 则 $Y=Y'\bmod 2\pi$， 则使 $X\to Y'$ 互信息最大与 $X\to Y$ 的互信息

最大是等效的。

而Y'的输出范围为：$(-a-\pi,a+\pi)$

$$
\begin{aligned}
I(X,Y') &= H(Y')-H(Y'/X)\\
&= \mathrm{H(Y')}-H(Z)\\
&\le \log(2\pi+2a)-\log(2a)\\
&= \log(\frac{\pi+a}{a})
\end{aligned}
$$

不等式成立的条件是 Y'为均匀分布，即 $p(y')=1/2(\pi+a) \quad y'\in(-\pi-a,\pi+a)$。

（1） $a\le\pi$ 时， $p(y)=\begin{cases}1/2(\pi+a) & 0\le y\le\pi-a \ \ , \ \ a+\pi\le y\le 2\pi\\ 1/(\pi+a) & \pi-a\le \mathrm{y}\le 2\pi\end{cases}$

即此时$C=\max(H(Y)-H(Y\mid X))$

$$
\begin{aligned}
&=\max(H(Y))-H(Z)\\
&=\int_0^{\pi-a}\frac{1}{2(\pi+a)}\log(2(\pi+a))dx+\int_{a+\pi}^{2\pi}\frac{1}{2(\pi+a)}\log(2(\pi+a))dx\\
&+\int_{\pi-a}^{\pi+a}\frac{1}{(\pi+a)}\log(2(\pi+a))dx+\int_{-a}^{a}\frac{1}{2a}\log(2a)dx\\
&=\frac{\pi-a}{\pi+a}\log 2(\pi+a)+\frac{2a}{a+\pi}\log(a+\pi)-\log 2a
\end{aligned}
$$

（2） $a>\pi$ 时， $p(y)=\begin{cases}3/2(\pi+a) & 0\le y\le a-\pi \ \ , \ \ 3\pi-a\le y\le 2\pi\\ 1/(\pi+a) & a\text{-}\pi\le y\le 3\pi-a\end{cases}$

则： $C=\max(H(Y)-H(Y\mid X))$

$$
\begin{aligned}
&=\max(H(Y))-H(Z)\\
&=\int_0^{a-\pi}\frac{3}{2(\pi+a)}\log(2(\pi+a)/3)dx+\int_{3\pi-a}^{2\pi}\frac{3}{2(\pi+a)}\log(2(\pi+a)/3)dx\\
&+\int_{a-\pi}^{3\pi-a}\frac{1}{(\pi+a)}\log(\pi+a)dx+\int_{-a}^{a}\frac{1}{2a}\log(2a)dx\\
&=\frac{3(a-\pi)}{a+\pi}\log\frac{2(a+\pi)}{3}+\frac{4\pi-2a}{a+\pi}\log(a+\pi)-\log 2a
\end{aligned}
$$

4.15 题的解答需要补充下面的知识

《信息论与编码》补充材料

熵功率不等式的证明

在信息论中熵功率不等式具有非常重要的应用，然后证明是非常繁琐。下面我们给出它的述叙和证明。

设 X，Y 为另均值独立随机变量，Z=X+Y，s_X^2，s_Y^2，s_Z^2 为 X，Y，Z 的功率（方差），相应的熵功率为 $\overline{s_X^2}$，$\overline{s_Y^2}$，$\overline{s_Z^2}$，则有如下不等式成立

$$\overline{s_X^2}+\overline{s_Y^2}\le\overline{s_Z^2}\le s_X^2+s_Y^2 \tag{1}$$

当 X，Y 为正态分布时，上式中的等号成立。

（随机变量 X 的熵功率定义为 $\overline{s_X^2}=\dfrac{1}{2pe}e^{2h(X)}$）

显然（1）式的右边不等式非常易证明：

因为 X 和 Y 独立，所以 $s_Z^2=s_X^2+s_Y^2$，由于熵功率不会大于功率，所以 $\overline{s_Z^2}\le s_Z^2=s_X^2+s_Y^2$。

所以关键是证明（1）式的左边。为此证明如下三条引理。

引理 1　令 $X_f=X+N(0,f)$，其中 X 的分布密度为 $p(x)$，$N(0,f)$ 为零均值，方差为 f 的正态分布随机变量，则

$$\frac{dh(X_f)}{df}=\frac{1}{2}\int\left(\frac{\partial p_f}{\partial x_f}\right)^2\frac{dx_f}{p_f} \tag{2}$$

其中 p_f 为 X_f 的概率分布密度。

[证明]　由于 $X_f=X+N(0,f)$

所以
$$p_f(x_f)=\frac{1}{\sqrt{2pf}}\int_{-\infty}^{\infty}p(x)\exp\left\{-\frac{(x_f-x)^2}{2f}\right\}dx \tag{3}$$

对（3）式中 f 求导得到

$$\frac{\partial p_f(x_f)}{\partial f}=\frac{1}{2}\frac{\partial^2 p_f(x_f)}{\partial x_f^2} \tag{4}$$

由（3）式可知

$$p_f(x_f)\le\frac{1}{\sqrt{2pf}}$$

X_f 的微分熵为

$$h(X_f)=-\int_{-\infty}^{\infty}p_f(x_f)\log p_f(x_f)dx_f$$

对 $h(X_f)$ **求关于** f **的导数得**

$$\begin{aligned}\frac{dh(X_f)}{df} &= -\int_{-\infty}^{\infty}\frac{\partial p_f}{\partial f}dx_f - \int_{-\infty}^{\infty}\frac{\partial p_f}{\partial f}\log p_f dx_f \\ &\overset{(a)}{=} 0 - \int_{-\infty}^{\infty}\frac{\partial p_f}{\partial f}\log p_f dx_f \\ &= -\frac{1}{2}\int_{-\infty}^{\infty}\frac{\partial^2 p_f(x_f)}{\partial x_f^2}\log p_f dx_f \\ &= -\frac{1}{2}\frac{\partial p_f}{\partial x_f}\log p_f\Bigg|_{-\infty}^{\infty} + \frac{1}{2}\int_{-\infty}^{\infty}\left(\frac{\partial p_f}{\partial x_f}\right)^2\frac{dx_f}{p_f} \\ &\overset{(b)}{=} 0 + \frac{1}{2}\int_{-\infty}^{\infty}\left(\frac{\partial p_f}{\partial x_f}\right)^2\frac{dx_f}{p_f}\end{aligned}$$

其中（a**）由于**
$$\begin{aligned}\int_{-\infty}^{\infty}\frac{\partial p_f}{\partial f}dx_f &= \frac{1}{2}\int_{-\infty}^{\infty}\frac{\partial^2 p_f(x_f)}{\partial^2 x_f}dx_f \\ &= \frac{1}{2}\frac{\partial p_f(x_f)}{\partial x_f}\Bigg|_{-\infty}^{\infty} \\ &=0\end{aligned}$$

（b**）因为**
$$\frac{\partial p_f}{\partial x_f} = -\frac{1}{\sqrt{2\pi f^3}}\int_{-\infty}^{\infty}p(x)(x_f - x)\exp\left\{-\frac{(x_f - x)^2}{2f}\right\}dx$$

由 Schwarz **不等式**

$$\left(\frac{\partial p_f}{\partial x_f}\right)^2 = \left[\int_{-\infty}^{\infty}\frac{\sqrt{p(x)}\exp\left\{-\frac{(x_f - x)^2}{4f}\right\}}{\sqrt[4]{2\pi f}}\cdot\frac{\sqrt{p(x)}(x_f - x)\exp\left\{-\frac{(x_f - x)^2}{4f}\right\}}{\sqrt[4]{2\pi f^5}}dx\right]^2$$

$$\le \frac{p_f}{\sqrt{2\pi f^5}}\int_{-\infty}^{\infty}p(x)\left(x_f - x\right)^2\exp\left\{-\frac{\left(x_f - x\right)^2}{2f}\right\}dx$$

$$\overset{(c)}{\le} \sqrt{\frac{2}{\pi f^3}}\cdot\frac{p_f}{e} \tag{5}$$

（c）由于$\left(x_f - x\right)^2 \exp\left\{-\frac{\left(x_f - x\right)^2}{2f}\right\} \le \frac{2f}{e}$

由（5）式，当$x_f \to \infty$，$p_f \to 0$，$\frac{\partial p_f}{\partial f} \sim 0\left(\sqrt{p_f}\right)$，所以$\frac{\partial p_f}{\partial x_f} \cdot \log p_f \to 0$；

从而证明了

$$\frac{dh\left(X_f\right)}{df} = \frac{1}{2}\int\left(\frac{\partial p_f}{\partial x_f}\right)^2 \cdot \frac{dx_f}{p_f} \qquad \text{Q.E.D}$$

对任何分布密度为$p(x)$的随机变量，记

$$J(X) = E\left[\left(\frac{p'}{p}\right)^2\right] = \int\left(\frac{\partial p}{\partial x}\right)^2 \cdot \frac{dx}{p} \tag{6}$$

则引理 1 可表示为$\frac{dH\left(X_f\right)}{df} = \frac{1}{2}J\left(X_f\right)$（7）

引理 2　X，Y 为独立随机变量，$Z = X + Y$，则，

$$\frac{1}{J(Z)} \ge \frac{1}{J(Y)} + \frac{1}{J(Y)} \tag{8}$$

［证明］设 X，Y 的分布密度分别为 $p(x)$，$q(y)$，则 Z 的分布密度为

$$r(z) = \int_{-\infty}^{\infty} p(x)q(z-x)dx$$

所以

$$r'(z) = \int_{-\infty}^{\infty} p(x)q'(z-x)dx$$

$$= \int_{-\infty}^{\infty} p'(x)q(z-x)dx$$

因而

$$\frac{r'(z)}{r(z)} = \int_{-\infty}^{\infty} \frac{p(x)\cdot q(z-x)}{r(z)} \cdot \frac{p'(x)}{p(x)}dx$$

$$= E\left\{\frac{p'(x)}{p(x)} \middle| z\right\}$$

其中$E\left\{\frac{p'(x)}{p(x)} \middle| z\right\}$表示在给定 z 条件下的条件期定。

同样可得到

$$\frac{r'(z)}{r(z)} = E\left\{\frac{q'(y)}{q(y)} \middle| z\right\}$$

于是 $$E\left\{a\frac{p'(x)}{p(x)}+b\frac{q'(y)}{q(y)}\Big|z\right\}=(a+b)\frac{r'(z)}{r(z)}$$

因而 $$(a+b)^2\left[\frac{r'(z)}{r(z)}\right]^2=\left[E\left\{a\frac{p'(x)}{p(x)}+b\frac{q'(y)}{q(y)}\Big|z\right\}\right]^2$$

$$\leq E\left\{\left[a\frac{p'(x)}{p(x)}+b\frac{q'(y)}{q(y)}\right]^2\Big|z\right\} \tag{9}$$

当 $$a\frac{p'(x)}{p(x)}+b\frac{q'(y)}{q(y)}=(a+b)\frac{r'(z)}{r(z)} \tag{10}$$

以概率 1 成立时，（9）式的等号才成立；而当 X，Y 为正态分布时，（10）式以概率 1 成立；所以当 X，Y 为正态分布时（9）式的等号成立。

（9）式二边对 Z 求平均，得到

$$(a+b)^2E\left\{\left[\frac{r'(z)}{r(z)}\right]^2\right\}\leq E\left\{\left[a\frac{p'(x)}{p(x)}+b\frac{q'(y)}{q(y)}\right]^2\right\}$$

$$=a^2E\left\{\left[\frac{p'(x)}{p(x)}\right]^2\right\}+b^2E\left\{\left[\frac{q'(y)}{q(y)}\right]^2\right\}$$

即 $$(a+b)J(Z)\leq a^2J(X)+bJ(Y)$$

令 $a=\dfrac{1}{J(X)}$，　$b=\dfrac{1}{J(Y)}$**，得到**

$$\frac{1}{J(Z)}\geq\frac{1}{J(X)}+\frac{1}{J(Y)}$$ **Q.E.D**

引理 3 令 $X_f=:X+N(0,f(t))$**，**$Y_g=Y+N(0,g(t))$**，其中** X、Y、$N(0,f(t))$ **和** $N(0,g(t))$ **为独立随机变量，**$N(0,g(t))$**，**$N(0,f(t))$ **为零均值，方差分别为** $g(t)$ **和** $f(t)$ **的正态分布，**

$$Z_l=X_f+Y_g=X+Y+N(0,l(t))$$

其中，$l(t)=f(t)+g(t)$**，记**

$$s(t)=\frac{\exp\{2h(X_f)\}+\exp\{2h(Y_g)\}}{\exp\{2h(Z_l)\}} \tag{11}$$

则 $s(0)\leq 1$

［证明］ 式（11）二边对 t 求导得

$$e^{2h(Z_l)}s'(t)=e^{2h(X_f)}\cdot f'(t)\cdot J(X_f)+e^{2h(Y_g)}\cdot g'(t)\cdot J(Y_g)-\left[e^{2h(X_f)}+e^{2h(Y_g)}\right]\cdot\left[f'(t)+g'(t)\right]\cdot J(Z_l)$$

代入不等式（8）**得**

$$e^{2h(Z_l)}\cdot s'(t)\geq\frac{\left[f'(t)\cdot J(X_f)-g'(t)\cdot J(Y_g)\right]\left[e^{2h(X_f)}\cdot J(X_f)-e^{2h(Y_g)}\cdot J(Y_g)\right]}{J(X_f)+J(Y_g)} \quad (12)$$

置 $f'(t)=e^{2h(X_f)}$ （13a）

$g'(t)=e^{2h(Y_g)}$ （13b）

则 $s'(t)\geq 0$**，所以** $s(+\infty)\geq s(0)$

下面证明 $s(+\infty)=1$。

由（13a）**和**（13b）**可知** $f(\infty)=g(\infty)=l(\infty)=\infty$

定义 $X_F=\dfrac{X_f}{\sqrt{f}}$ **则**

$$P_F(x_F)=\frac{1}{\sqrt{2p}}\int_{-\infty}^{\infty}\sqrt{f}\,p\left(\sqrt{f}u\right)\exp\left\{-\frac{1}{2}(x_F-u)^2\right\}du$$

$$h(X_F)=h(X_f)-\frac{1}{2}\log f$$

当 $f\to\infty$ **时，** $\sqrt{f}\,f\left(\sqrt{f}u\right)\to d(u)$，

所以当 $f\to\infty$ **时，** $P_F(x_F)\to\exp\left\{-\dfrac{x_F^2}{2}\right\}/\sqrt{2p}$

于是 $h(X_F)\to\dfrac{1}{2}\log 2pe$

所以 $h(X_f)-\dfrac{1}{2}\log 2pef\to 0$ （14a）

同样 $h(Y_g)-\dfrac{1}{2}\log 2peg\to 0$ （14b）

$$h(Z_e)-\frac{1}{2}\log 2pe(f+g)\to 0 \qquad (14c)$$

由于当 $t\to\infty$ 时，$f(t)$，$g(t)$，$l(t)\to\infty$

由（11）和（14）式

$$s(\infty)=1$$

所以　　$s(0)\le s(\infty)=1$　　**Q.E.D**

由引理 3 ，若设 $f(0)=g(0)=0$，则

$$s(0)=\frac{e^{2h(X)}+e^{2h(Y)}}{e^{2h(Z)}}\le 1$$

所以　　$e^{2h(Z)}\ge e^{2h(X)}+e^{2h(Y)}$

因此　　$\overline{s}_Z^{\,2}\ge \overline{s}_X^{\,2}+\overline{s}_Y^{\,2}$

当 X，Y 是正态分布时不等式中等号成立。

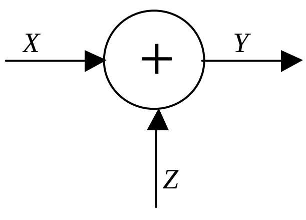
利用熵功率不等式可以证明习题 4.15。

4.15 考虑如图所示信道，对于信号和噪声功率有如下的限制：

$$EX=0, EX^2=P, EZ=0, EZ^2=N$$

如果记 X^* 为高斯随机变量 N$(0,P)$，Z^* 为高斯随机变量 N$(0,N)$，试证明：对 X，Z 的任何其它分布有

$$I(X;X+Z^*)\le I(X^*;X^*+Z^*)\le I(X^*;X^*+Z)$$



习题 4.15 图

［证明］ 当 Z=Z*

$$I(X;X+Z^*)=h(X+Z^*)-h(X+Z^*|X)$$
$$=h(X+Z^*)-h(Z^*)$$

$$=h\left(X+Z^*\right)-\frac{1}{2}\log\left(2peN\right)$$

$$E\left[\left(X+Z^*\right)^2\right]=P+N`，$$

在所有功率为 $P+N$ **的随机变量中，当**$X+Z^*$**为正态分布时其熵为最大，而当** $X\sim N$（0，P）**时，**$X+Z^*$**为正态，即** X^*+Z^***，所以**

$$I\left(X;X+Z^*\right)\le h\left(X^*+Z^*\right)-\frac{1}{2}\log\left(2peN\right)$$

$$=I\left(X^*;X^*+Z^*\right)$$

当 $X=X^*$**时，由熵功率不等式**

$$e^{2\cdot I(X^*;\,X^*+Z)}=\frac{e^{2h(X^*+Z)}}{e^{2h(Z)}}\ge\frac{e^{2h(X^*)}+e^{2h(Z)}}{e^{2h(Z)}}$$

$$=1+\frac{e^{2h(X^*)}}{e^{2h(Z)}}$$

$$\ge 1+\frac{e^{2h(x^*)}}{e^{2h(Z^*)}}$$

$$=e^{2\cdot I(X^*;\,X^*+Z^*)}$$

当 Z **为正态变量** Z^***时，上式中二个不等式中的等号均成立，由熵功率不等式得**

$$I\left(X^*;X^*+Z^*\right)\le I\left(X^*;X^*+Z\right)\qquad\textbf{Q.E.D}$$