3.1.2 单符号离散信道-常用到的概率关系

1) 先验概率 $p(a_i)=p(X=a_i) \qquad i=1,\cdots,r$

2) 联合概率 $p(a_ib_j)=p(X=a_i,Y=b_j)$

并且有 $p(a_ib_j)=p(a_i)p(b_j|a_i)=p(b_j)p(a_i|b_j)$

3) 前向概率（即信道传递概率）

$$p(b_j|a_i)=p(Y=b_j|X=a_i)$$

4) 后向概率（又称后验概率）

$$p(a_i|b_j)=p(X=a_i|Y=b_j)$$

3.1.2 单符号离散信道-常用到的概率关系

$$p(a_i|b_j)=\frac{p(a_ib_j)}{p(b_j)}=\frac{p(a_i)p(b_j|a_i)}{\sum_{i=1}^{r}p(a_ib_j)}=\frac{p(a_i)p(b_j|a_i)}{\sum_{i=1}^{r}p(a_i)p(b_j|a_i)}$$

5) 输出符号概率

$$p(b_j)=p(Y=b_j) \qquad j=1,\cdots,s$$

$$p(b_j)=\sum_{i=1}^{r}p(a_ib_j)=\sum_{i=1}^{r}p(a_i)p(b_j|a_i)$$

用矩阵形式表示为

$$\begin{bmatrix}p(b_1) & p(b_2) & \cdots & p(b_s)\end{bmatrix}=\begin{bmatrix}p(a_1) & p(a_2) & \cdots & p(a_r)\end{bmatrix}\cdot \boldsymbol{P}$$

$$\Leftrightarrow \quad \mathbf{P}_Y=\mathbf{P}_X\mathbf{P}_{Y|X}$$

3.2.1互信息量的定义

- 事件y_j所给出关于事件x_i的信息，即

$$I(x_i;y_j)=\underbrace{I(x_i)}_{\text{先验不定度}}-\underbrace{I(x_i|y_j)}_{\text{后验不定度}}=\log\frac{p(x_i|y_j)}{p(x_i)}$$

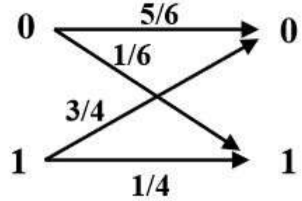
  - $I(x_i)$：收到y_j前对于信道输入符号x_i存在的不确定度；
  - $I(x_i|y_j)$：收到y_j后对x_i仍然存在的不确定度
- 互信息有两方面的含义
  - 表示事件y_j出现前后关于事件x_i的不确定性减少的量
  - 事件y_j出现以后信宿获得的关于事件x_i的信息量

3.2. 1互信息量的定义

例3.6：

一离散无记忆信源，其概率空间为 $\begin{bmatrix} X \\ P(X) \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ 0.6 & 0.4 \end{bmatrix}$

信源发出的符号通过一干扰信道，信道输出端的接收符号集为$Y=[0,1]$，信道传递概率如下所示：



（1）信源符号0和1分别含有的自信息量；

$I(x=0)$和$I(x=1)$?

（2）收到y_j后，获得的关于x_i的信息量；

$I(x=0;y=0)$, $I(x=0;y=1)$, $I(x=1;y=0)$, $I(x=1;y=1)$

（3）X和Y的熵。

$H(X)$, $H(Y)$?

3.2. 2互信息的性质

1. 互信息的互易性

用公式表示为 $I(x_i;y_j)=I(y_j;x_i)$
互信息的对称性表明：

从y_j得到的关于x_i的信息量 $I(x_i;y_j)$ 与从x_i得到的关于y_j的信息量 $I(y_j;x_i)$ 是一样的，只是观察的角度不同而已。

3.2. 2互信息的性质- 3.互信息可正可负

$$I(x_i;y_j)=\log\frac{1}{p(x_i)}-\log\frac{1}{p(x_i|y_j)}=I(x_i)-I(x_i|y_j)$$

- $p(x_i|y_j)=1$　有$I(x_i;y_j)=I(x_i)$，表明收到 y_j 可完全消除对于 x_i 的不确定度，从而获取 x_i 的全部信息量
- $p(x_i)<p(x_i|y_j)<1$ 有$I(x_i;y_j)>0$，表明收到 y_j 后判断信源发 x_i 的可能性大于收到 y_j 前判断信源发 x_i 的可能性，后验概率相对先验概率的比值越大，互信息量越大
- $p(x_i|y_j)=p(x_i)$　有$I(x_i;y_j)=0$，表明收到 y_j 后并没有获得关于 x_i 的信息量
- $0<p(x_i|y_j)<p(x_i)$ 有$I(x_i;y_j)<0$，表明收到 y_j 非但没有减少，反而增加了对信源发 x_i 的不确定度

3.2. 2互信息的性质

2. 互信息量可为0

- x_i 和 y_j相互独立,互信息为0

$$I(x_i;y_j)=\log\frac{1}{p(x_i)}-\log\frac{1}{\underbrace{p(x_i|y_j)}_{=p(x_i)}}=I(x_i)-I(x_i|y_j)$$

$$=0$$

- 表明收到 y_j 后并没有消除任何不确定度，因此不能获得关于 x_i的信息量

3.2. 2互信息的性质

4.任何两个事件之间的互信息不可能大于其中任一事件的自信息

$$I(x_i;y_j)=\log\frac{p(x_i|y_j)}{p(x_i)}\le\log\frac{1}{p(x_i)}=I(x_i)$$

$$I(y_j;x_i)=\log\frac{p(y_j|x_i)}{p(y_j)}\le\log\frac{1}{p(y_j)}=I(y_j)$$

自信息量是确定该事件出现的所必需的信息量，也是任何其他事件所能提供的关于该事件的最大信息量

3.2. 3 条件互信息

- 联合集*XYZ*中，在给定事件y_j的条件下，x_i与z_k之间的互信息定义为***条件互信息***:

$$I(x_i;z_k\mid y_j)=\log\frac{p(x_i\mid y_jz_k)}{p(x_i\mid y_j)}$$

条件互信息和互信息的区别仅在于先验概率和后验概率都是在<u>某一特定条件下</u>的取值

3.2.4 联合互信息

- 联合集*XYZ*中，x_i与y_jz_k的互信息量

$$I(x_i;y_jz_k)=\log\frac{p(x_i\mid y_jz_k)}{p(x_i)}=\log\left[\frac{p(x_i\mid y_jz_k)}{p(x_i)}\bullet\frac{p(x_i\mid y_j)}{p(x_i\mid y_j)}\right]$$

$$=I(x_i;y_j)+I(x_i;z_k\mid y_j)$$

- 这说明 y_jz_k 联合发生后所提供的有关 x_i 的信息量 $I(x_i;y_jz_k)$，等于 y_j 给出关于 x_i 的信息量$I(x_i;y_j)$加上y_j已知条件下 z_k 多给出关于 x_i 的信息量$I(x_i;z_k|y_j)$。
- 以上关系可推广到任意维空间，在<u>网络信息论和概率图理论及应用</u>中这些关系十分有用

3.2.5平均互信息的定义

- 是互信息量 $I(x_i;y_j)$在<u>联合概率空间$P(XY)$</u>中的统计平均值，被定义为随机变量 X 和随机变量 Y 之间的平均互信息，是两个概率空间 X 和 Y 之间的平均互信息，用公式表示为

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)I(x_i;y_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

$$=\sum_{j}p(y_j)\sum_{i}p(x_i|y_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

$I(X;y_j)$

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{1}{p(x_i)}-\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{1}{p(x_i|y_j)}$$

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

条件熵表示给定随机变量Y后，对随机变量X仍然存在的不确定度，也称疑义度

$$=H(X)+H(Y)-H(XY)$$

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

条件熵表示给定随机变量X后，对随机变量Y仍然存在的不确定度，完全由于信道中的噪声引起，又称噪声熵（散布度）

3.2.5平均互信息的定义

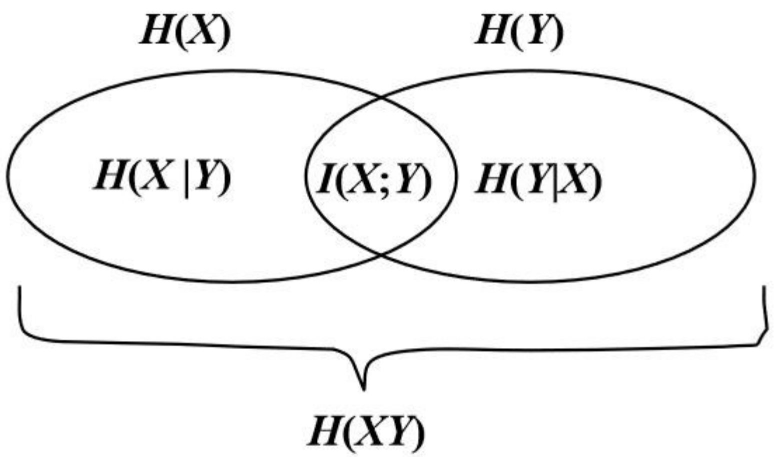
- <u>信道疑义度</u>表示接收端收到信道输出的一个符号之后对信道输入的符号仍然存在的平均不确定性。

$$H(X|Y)=E_Y[H(X|y_j)]=-\sum_{i=1}^{r}\sum_{j=1}^{s}p(x_iy_j)\log p(x_i|y_j)$$

✓ 理想信道，$H(X|Y)=0$。

✓ 一般情况下，$H(X|Y)<H(X)$。

✓ 当$H(X|Y)=H(X)$时，表示接收到输出变量Y后关于输入变量X的平均不确定性一点也没有减少。



维拉图－一种重要的分析工具

表1. 随机变量X、Y之间有依赖关系时

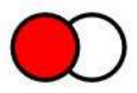
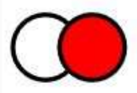
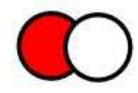
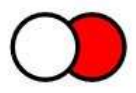
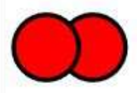
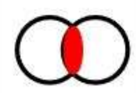
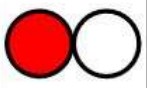
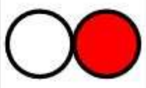
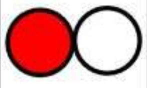
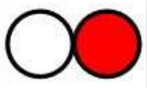
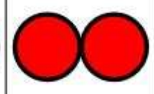
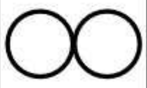
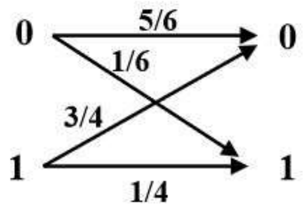
图示						
符号	$H(X)$	$H(Y)$	$H(X\|Y)$	$H(Y\|X)$	$H(XY)$	$I(X;Y)$

表2. 随机变量X、Y之间相互独立时

图示						
符号	$H(X)$	$H(Y)$	$H(X\|Y)=H(X)$	$H(Y\|X)=H(Y)$	$H(XY)=H(X)+H(Y)$	$I(X;Y)=0$

例：一离散无记忆信源，其概率空间为 $\begin{bmatrix}X\\P(X)\end{bmatrix}=\begin{bmatrix}0 & 1\\0.6 & 0.4\end{bmatrix}$

信源发出的符号通过一干扰信道，信道输出端的接收符号集为Y=[0, 1]，信道传递概率如下所示：



（1）收到y_j后，获得的关于x_i的信息量；

$I(x=0;y=0)$, $I(x=0;y=1)$, $I(x=1;y=0)$, $I(x=1;y=1)$

（2）求$I(X;Y)$

3种不同的解法

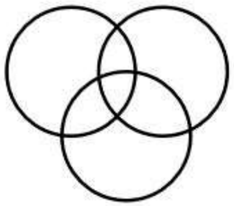
3.2.6平均联合互信息

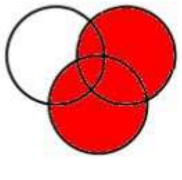
- 将联合互信息$I(x;yz)$在概率空间XYZ中求统计平均

$$I(X;YZ)=E[I(x;yz)]=\sum_X\sum_Y\sum_Z p(xyz)\log\frac{p(x\mid yz)}{p(x)}$$

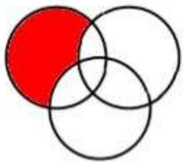
$$=H(X)-H(X\mid YZ)$$

维拉图表示？

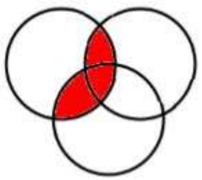




$H[YZ]$



$H[X/YZ]$



$H(X)-H(X|YZ)$

3.2.7平均条件互信息

- 将条件互信息$I(x;z|y)$在概率空间XYZ中求统计平均

$$I(X;Z\mid Y)=E[I(x;z\mid y)]=\sum_X\sum_Y\sum_Z p(xyz)\log\frac{p(x\mid yz)}{p(x\mid y)}$$

$$=\sum_X\sum_Y\sum_Z p(xyz)\log\frac{p(x\mid yz)}{p(x)}\cdot\frac{p(x)}{p(x\mid y)}$$

$$=I(X;YZ)-I(X;Y)$$

$$=H(X)-H(X\mid YZ)-(H(X)-H(X\mid Y))$$

$$=H(X\mid Y)-H(X\mid YZ)\geq 0$$

表示收到Y后再收到Z所多获得的关于X的信息量

维拉图表示？

与三变量有关的平均互信息分析 – **维拉图并不是万能的！**

$I(X;Y)$? $I(X;Y|Z)$

1. $Y=X$ $Z=Y$
   $I(X;Y)=H(X)-H(X|Y)=H(X)$
   $I(X;Y|Z)=H(X|Z)-H(X|YZ)=0$

2. X和Y相互独立 $Z=X+Y$
   $I(X;Y)=H(X)-H(X|Y)=0$
   $I(X;Y|Z)=H(X|Z)-H(X|YZ)=H(X|Z)$

事实上，不存在度量三个随机变量所共有的互信息概念
根据维拉图，可定义
$I(X;Y;Z)=I(X;Y)-I(X;Y|Z)$ 不一定非负

3.3 平均互信息的性质

1.非负性 $I(X;Y)\ge 0$

证明 利用 $\ln x \le x-1 \quad (x>0)$

$$-I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_i)}{p(x_i\mid y_j)}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(y_j)p(x_i\mid y_j)\log\frac{p(x_i)}{p(x_i\mid y_j)}$$

$$\le(\log_2 e)\sum_{i=1}^{n}\sum_{j=1}^{m}p(y_j)p(x_i\mid y_j)\left(\frac{p(x_i)}{p(x_i\mid y_j)}-1\right)$$

$$=(\log_2 e)\sum_{i=1}^{n}\sum_{j=1}^{m}\left(p(x_i)p(y_j)-p(x_iy_j)\right)=0$$

也可用詹森不等式

$$\le\sum_{j=1}^{m}p(y_j)\log\sum_{i=1}^{n}p(x_i/y_j)\frac{p(x_i)}{p(x_i/y_j)}$$

$$=\sum_{j=1}^{m}p(y_j)\log\sum_{i=1}^{n}p(x_i)=\sum_{j=1}^{m}p(y_j)\log 1=0$$

- 当X和Y统计独立时，$I(X;Y)=0$
- 平均互信息量不是从两个具体消息出发，而是从随机变量X和Y的整体角度出发，在平均意义上观察问题，所以平均互信息不会出现负值。
- 因此，给定随机变量Y后，一般说来总能消除一部分关于X的不确定性。

3.3 平均互信息的性质

2. 对称性 $I(X;Y)=I(Y;X)$

证明 $$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_i\mid y_j)}{p(x_i)}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_iy_j)}{p(x_i)p(y_j)}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(y_j\mid x_i)}{p(y_j)}$$

$$=I(Y;X)$$

对称性说明，从Y中获得关于X的信息量等于从X中获得关于Y的信息量

3.3 平均互信息的性质

3. 极值性 $I(X;Y)=H(X)-H(X|Y)$

$$I(X;Y)\le H(X)$$

说明：接收者通过信道获得的信息量不可能超过信源本身固有的信息量。只有当疑义度$H(X|Y)=0$时，接收到Y时才能获得全部关于X的信息量。

一般情况下，平均互信息在0和$H(X)$值之间。

- 平均互信息是信道输入概率分布和信道传递概率的凸函数。

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)I(x_i;y_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

可根据输入概率分布和信道传递概率计算

- 两个重要定理
  - 对于固定信道，即$p(y_j|x_i)$不变的前提下，$I(X;Y)$是输入信源概率分布的上凸函数。
  - 对于固定信源，即$p(x)$分布不变的前提下，$I(X;Y)$是信道传递概率$p(y_j|x_i)$的下凸函数。

3.3 平均互信息的上凸性定理

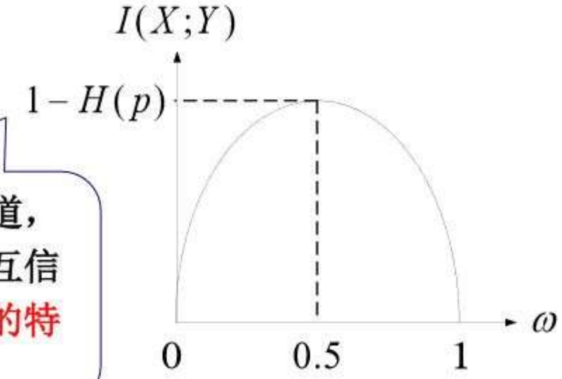
- 对于固定的信道，平均互信息是信源概率分布 $p(x)$ 的上凸函数。$f[\alpha X_1+(1-\alpha)X_2]\ge \alpha f(X_1)+(1-\alpha)f(X_2)$

***$I[p(x_i)]$*是输入信源的概率分布*$p(x_i)$*的"∩"型凸函数。**

$$I(X;Y)=H(Y)-H(Y|X)$$
$$=H(Y)-\sum_{i=1}^{2}\sum_{j=1}^{2}p(x_i)p(y_j|x_i)\log p(y_j|x_i)$$
$$=H(Y)-\sum_{i=1}^{2}p(x_i)\{-[\bar p\log\bar p+p\log p]\}$$
$$\boxed{\boldsymbol{P}=\begin{bmatrix}\bar p & p\\ p & \bar p\end{bmatrix}}$$
$$=H(Y)-\{-[\bar p\log\bar p+p\log p]\}=H(Y)-H(p)$$
$$p(y_1=0)=p(x=0)p(0|0)+p(x=1)p(0|1)=w\bar p+\bar w p$$
$$p(y_2=1)=p(x=0)p(1|0)+p(x=1)p(1|1)=wp+\bar w\bar p$$
$$H(Y)=(w\bar p+\bar w p)\log\frac{1}{w\bar p+\bar w p}+(wp+\bar w\bar p)\log\frac{1}{wp+\bar w\bar p}$$
$$=H(w\bar p+\bar w p)$$
$$\boxed{\therefore I(X;Y)=H(w\bar p+\bar w p)-H(p)}$$

1) $I(X;Y)=H(Y)-H(Y|X)=H(\omega\bar p+\bar\omega p)-H(p)$

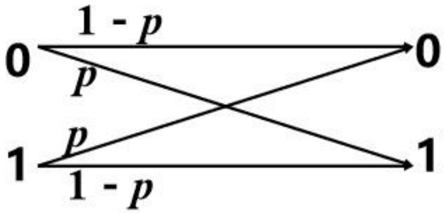
2) **固定信道，当** $\omega=\bar\omega=\frac{1}{2}$ **时，** $H(\omega\bar p+\bar\omega p)=H(\frac{1}{2})=1$ **互信息取得最大值。**



物理意义：对某一个确定信道，存在一种信源分布，使平均互信息最大。最大值由信道本身的特性决定。

例3.8：二进制对称信道，输入X概率分布,即信源概率空间为：
$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \omega & 1-\omega \end{bmatrix}$$

信道如图：

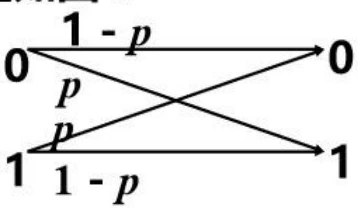


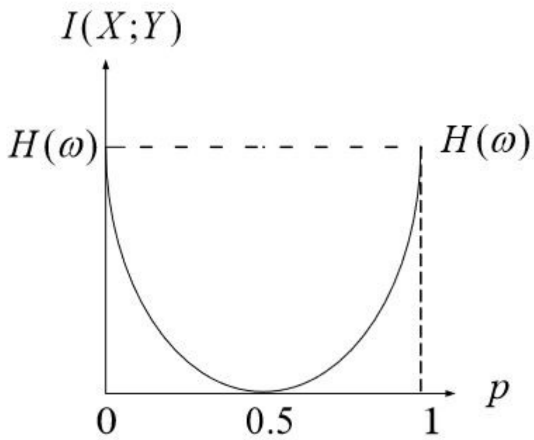
试分析平均互信息与信源概率分布的关系

例3.9：二进制对称信道，输入X 概率分布,即信源概率空间为：

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \omega & 1-\omega \end{bmatrix}$$

信道如图：





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

关于下凸性定理的说明

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

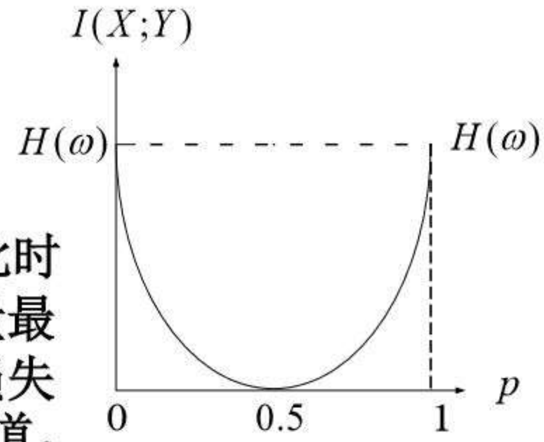
p＝0时，无噪信道

$$I(X;Y) = H(w);$$

p＝1时，无噪信道：

$$I(X;Y) = H(w);$$

p＝1/2时： $I(X;Y) = 0$，此时在信道输出端获得信息量最小，即信源的信息全部损失在信道中——>最差的信道。



<u>物理意义</u>：每种信源都存在一种对应的最差信道，此信道的干扰最大，输出端获得的信息量最小

3.3 平均互信息的特性

－各种熵、信道疑义度、平均互信息之间的关系

- 联合熵与信源熵
  - $H(XY) \le H(X) + H(Y)$ - 独立时，等号成立
- 联合熵、信源熵、条件熵
  - $H(XY) = H(X|Y) + H(Y) = H(Y|X) + H(X)$
- 平均互信息与熵
  - $I(X;Y) = H(X) - H(X|Y)$
  - $I(X;Y) = H(X) + H(Y) - H(XY)$
  - $I(X;Y) = I(Y;X) \ge 0$
  - $I(X;X) = H(X)$

3.4 信道容量及其一般计算方法

信道容量的定义

- <u>信息传输率 R</u>：信道中平均每个符号所传送的信息量。
- 平均互信息 $I(X;Y)$是接收到符号Y后平均获得的关于X的信息量。所以

$$R = I(X;Y) = H(X) - H(X|Y) \quad (\text{比特/符号})$$

- 设平均传输一个符号需要t秒，则信道每秒钟平均传输的信息量为信息传输速率 R_t:

$$R_t = \frac{1}{t} I(X;Y) \quad （比特/秒）$$

3.4 信道容量的定义

- 在信道确定的情况下，$R = I(X;Y)$是信源概率分布$P(X)$的上凸函数。因此，必然存在一种信源概率分布使信息传输率$R = I(X;Y)$最大。定义这个最大的信息传输率为<u>信道容量</u>：

$$C = \max_{P(X)} \{I(X;Y)\} \quad （比特/符号）$$

相应的输入概率分布被称为<u>最佳输入分布</u>。

- **信道容量：**
  - **与信源的概率分布无关；**
  - **是完全描述信道特性的参量；**
  - **是信道能够传输的最大信息量。**

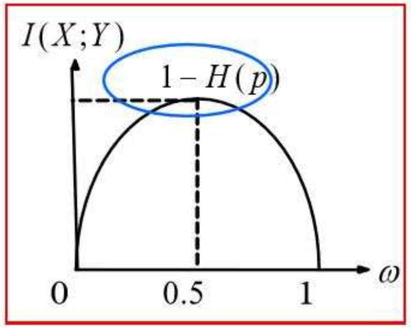
$$\boxed{C_t = \frac{1}{t} \max_{P(X)} \{I(X;Y)\} \quad （比特/秒）}$$

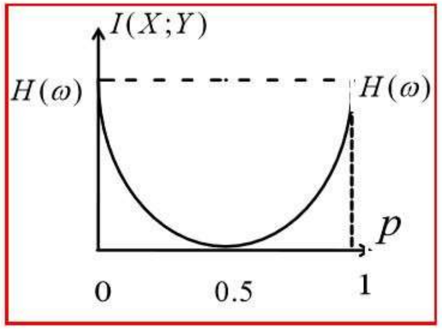
3.4 信道容量的定义

- **二元对称信道的信道容量**

$$\boldsymbol{P} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

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





3.4 信道容量的求解

1.离散无损信道 输出*Y*与输入*X*之间有着确定关系

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

2.离散无噪信道 多个输入对应一个输出

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

3.离散无噪无损信道 输入和输出一一对应

无噪无损信道容量 C $=\log r=\log s$

离散对称信道定理

<u>对称信道</u> 当信道输入概率分布为等概的情况下达到信道容量：

$$C=\log s-H(p_1',p_2',\cdots,p_s')$$

$$\mathbf{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{3} & \frac{1}{6}\\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2}\end{bmatrix}$$

$$\begin{aligned}C&=\log s-H(p_1',p_2',\cdots,p_s')\\&=\log 3-H(\frac{1}{2},\frac{1}{3},\frac{1}{6})\\&=0.126\text{比特/符号}\end{aligned}$$

离散强对称信道

- **均匀信道的信道容量**

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

$$\begin{aligned} C &= \log s - H(p_1', p_2', \cdots, p_s') \\ &= \log r - H(\bar{p}, \frac{p}{r-1}, \cdots, \frac{p}{r-1}) \\ &= \log r + \bar{p}\log\bar{p} + \frac{p}{r-1}\log\frac{p}{r-1} + \cdots + \frac{p}{r-1}\log\frac{p}{r-1} \\ &= \log r + \bar{p}\log\bar{p} + p\log\frac{p}{r-1} \\ &= \log r + \bar{p}\log\bar{p} + p\log p - p\log(r-1) \\ &= \underline{\log r - H(p) - p\log(r-1)} \end{aligned}$$

离散准对称信道容量的求解

- 例3.11 $P=\begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$ $P_1=\begin{bmatrix} 0.8 & 0.1 \\ 0.1 & 0.8 \end{bmatrix}$ $P_2=\begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$

当输入等概分布时到达信道容量，此时的输出概率分布为：

$$[1/2 \quad 1/2]\begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} = [0.45 \quad 0.1 \quad 0.45]$$

$$\begin{aligned} C &= H(Y) - H(0.8, 0.1, 0.1) \\ &= H(0.45, 0.1, 0.45) - H(0.8, 0.1, 0.1) \end{aligned}$$

或者根据公式：

$N_1=0.9, M_1=0.9$
$N_2=0.1, M_2=0.2$

$$\begin{aligned} C &= \log 2 - H(0.8, 0.1, 0.1) - \sum_{k=1}^{2} N_k \log M_k \\ &= 1 - (0.9\log 0.9 + 0.1\log 0.2) - H(0.8, 0.1, 0.1) \end{aligned}$$

一般信道容量的求解

①由$\sum_{j=1}^{s} p(y_j|x_i)\beta_j = \sum_{j=1}^{s} p(y_j|x_i)\log p(y_j|x_i)$，求$\beta_j$；

②由$C=\log\left(\sum_{j=1}^{s} 2^{\beta_j}\right)$，求$C$；

③由$p(y_j)=2^{\beta_j - C}$，求$p(y_j)$；

④由$p(y_j)=\sum_{i=1}^{r} p(x_i)p(y_j|x_i)$,求$p(x_i)$。

- **例3.13信道矩阵如右所示，求信道容量**

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

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

$$\boxed{\sum_j p(b_j|a_i)\beta_j=\sum_j p(b_j|a_i)\log p(b_j|a_i) \quad i=1,2,\cdots,r}$$

$$\beta_2=\beta_3=0,\ \ \beta_1=\beta_4=-2$$

$$C=\log\sum_j 2^{\beta_j}=\log(2^{-2}+2^0+2^0+2^{-2})=\log 5-1 \text{ 比特/符号}$$

$$p(b_j)=2^{\beta_j - C} \qquad C=\log 5-1$$

$$\beta_2=\beta_3=0,\ \ \beta_1=\beta_4=-2$$

$$p(b_1)=p(b_4)=2^{-2-\log 5+1}=\frac{1}{10}$$

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

$$\boxed{p(b_j)=\sum_i p(a_i)p(b_j|a_i)}$$

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

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

3.4.4 信道容量定理

$$\boxed{\begin{aligned} I(a_i;Y) &= C \quad p(a_i)\neq 0 \\ I(a_i;Y) &\le C \quad p(a_i)=0 \end{aligned}}$$

$$\boxed{I(a_i;Y)=\sum_j p(b_j|a_i)\log\frac{p(b_j|a_i)}{p(b_j)}}$$

$$\frac{\partial I(X;Y)}{\partial p(a_i)}=\sum_{j=1}^{s} p(b_j|a_i)\log\frac{p(b_j|a_i)}{p(b_j)}-\log e$$

$$=I(a_i;Y)-\log e=C-\log e=\lambda+\log e-\log e=\lambda \qquad p(a_i)\neq 0$$

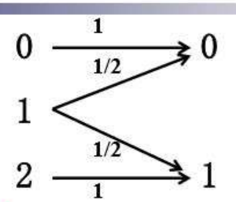
$$\frac{\partial I(X;Y)}{\partial p(a_i)}\le\lambda \qquad p(a_i)=0$$

充分性：假设有一个输入分布满足上式，则该输入分布一定可以使平均互信息达到最大值。

必要性：假设有一个输入分布可以使平均互信息达到最大值，则该输入分布一定满足上式。

3.4.4 信道容量定理- 例3.14

离散信道如图所示，求信道容量



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

信道传递矩阵

输入1的概率分布为0时，该信道成为无噪信道；当$p(0)=p(2)=1/2$时，信息传输达到信道容量log 2，此时：

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

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

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

可见，上述结果满足满足充要条件：

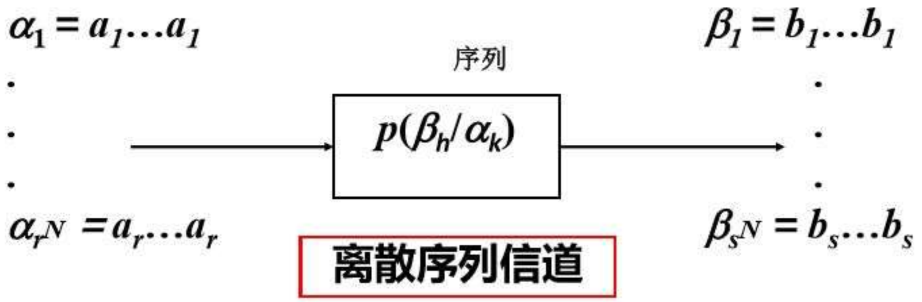
$$\begin{cases} I(a_i;y)=\log 2\text{时}, & p(a_i)\neq 0 \\ I(a_i;y)=0\le\log 2\text{时}, & p(a_i)=0 \end{cases}$$

达到该信道容量的输入概率就是前面假设的输入分布，即：$p(0)=p(2)=1/2$，$p(1)=0$

利用该定理可证明：实现准对称离散无记忆信道信道容量的输入符号分布为等概分布。

离散序列信道

- 一般离散信道输入和输出是一个随机变量序列
- 信道的输入或输出均取值于同一输入或输出符号集合



$$[P^N] = \begin{bmatrix} p(\beta_1 \mid \alpha_1) & p(\beta_2 \mid \alpha_1) & \cdots & p(\beta_{s^N} \mid \alpha_1) \\ p(\beta_1 \mid \alpha_2) & p(\beta_2 \mid \alpha_2) & \cdots & p(\beta_{s^N} \mid \alpha_2) \\ \cdots\cdots & & & \\ p(\beta_1 \mid \alpha_{r^N}) & p(\beta_2 \mid \alpha_{r^N}) & \cdots & p(\beta_{s^N} \mid \alpha_{r^N}) \end{bmatrix}$$

$$p_{ij} = p(\beta_j \mid \alpha_i) \quad p_{ij} \geq 0; \qquad \sum_{j=1}^{s^N} p_{ij} = 1 \quad (i = 1,2,\ldots,r^N)$$

信道矩阵中，各传递概率一般通过测量的方法给定。而对于离散无记忆信道，概率值可通过下面方法得到：

$$p(\beta_j \mid \alpha_i) = p(b_{j1} b_{j2} \ldots b_{jN} \mid a_{i1} a_{i2} \ldots a_{iN}) = \prod_{k=1}^{N} p(b_{jk} \mid a_{ik})$$

离散序列信道的平均互信息

- 离散序列信道的平均互信息定义：

$$I(\mathrm{X};\mathrm{Y}) = \sum_{X^N Y^N} p(\alpha_k \beta_h) \log \frac{p(\alpha_k \mid \beta_h)}{p(\alpha_k)}$$

$$= \sum_{X^N Y^N} p(\alpha_k \beta_h) \log \frac{p(\beta_h \mid \alpha_k)}{p(\beta_h)}$$

$$(k = 1,2,\cdots\cdots,r^N), \qquad (h = 1,2,\cdots\cdots,s^N)$$

对于一般离散信道

若信道输入/出分别是N长序列X和Y，且信道无记忆，则存在 $I(\mathrm{X};\mathrm{Y}) \leq \sum_{k=1}^{N} I(X_k;Y_k)$

- 若信源为无记忆信源时，等号成立

物理意义：

离散无记忆信道的N次扩展信道的平均互信息不会超过N个随机变量x_1，x_2，…，x_N单独通过离散无记忆信道的平均互信息之和；

仅当信源为离散无记忆信源时两者才相等

当信源无记忆时，等号成立，即 $I(\mathbf{X};\mathbf{Y})=\sum_{k=1}^{N}I(X_k;Y_k)$

$$p(\beta_j)=\sum_{i_1=1}^{r}\cdots\sum_{i_N=1}^{r}\cdot p(a_{i_1}\ldots a_{i_N})p(b_{j_1}\ldots b_{j_N}\mid a_{i_1}\ldots a_{i_N})$$

$$=\sum_{i_1=1}^{r}p(a_{i_1})p(b_{j_1}\mid a_{i_1})\cdots\sum_{i_N=1}^{r}p(a_{i_N})p(b_{j_N}\mid a_{i_N})$$

公式

$$=p(b_{j_1})\cdots p(b_{j_N})$$

$$I(\mathbf{X};\mathbf{Y})-\sum_{k=1}^{N}I(X_k;Y_k)=H(\mathbf{Y})-H(Y^N)\le 0$$

$H(\mathbf{Y})$——信源有记忆时，$\mathbf{Y}$的熵；

$H(Y^N)$——信源无记忆时，$\mathbf{Y}$的熵；

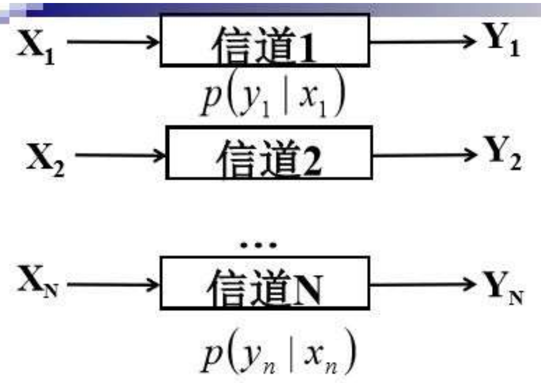
即，对同样信道
信源有记忆时的输出熵
≤
信源无记忆时的输出熵

可映证信源分析的结论！

独立并联信道及其信道容量

当信源、信道都是无记忆时，等号成立，相当于N个独立信道并联的情况

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



特点：每个信道的输出仅与本信道的输入有关

当信道矩阵相同，且输入变量独立时，则

$$C_{P^N}=NC_P$$

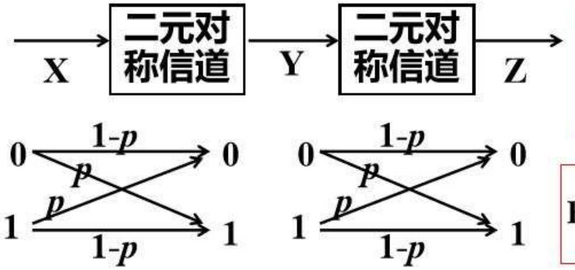
$$I(\mathbf{X};\mathbf{Y})=I(X_1X_2\cdots X_N;Y_1Y_2\cdots Y_N)\le\sum_{i=1}^{N}I(X_i;Y_i)$$

$$C=\max_{p(x_1,x_2,\cdots,x_N)}\{I(X_1X_2\cdots X_N;Y_1Y_2\cdots Y_N)\}\le\sum_{i=1}^{N}C_i$$

当输入变量相互独立且各输入变量的概率分布达到最佳分布时，等号成立

$$\boxed{C_i=\max_{p(x_i)}\{I(X_i;Y_i)\}}$$

级联信道



$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

$$\mathbf{P}_1 = \mathbf{P}_2 = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

$$\mathbf{P} = \mathbf{P}_1\mathbf{P}_2 = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

$$= \boxed{\begin{bmatrix} (1-p)^2 + p^2 & 2p(1-p) \\ 2p(1-p) & (1-p)^2 + p^2 \end{bmatrix}}$$

- **平均互信息**

通过第一个信道 $I(X;Y) = 1 - H(p)$

通过两个信道 $I(X;Z) = 1 - H[2p(1-p)]$

$$I(X;W) = 1 - H\left[3p(1-p)^2 + p^2\right]$$

说明

1. 信道串联后增加信息损失

2. 串联级数越多，损失越大

匹配：当信源与信道连接时，若信息传输率达到了信道容量，则称此信源与信道达到匹配。否则认为信道有剩余。

信道剩余度 = C - $I(X;Y)$

相对剩余度 = $[C$ - $I(X;Y)] / C$

= 1 - $I(X;Y) / C$

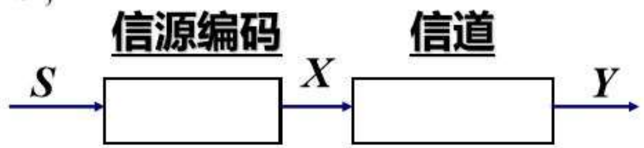
无损信道的相对剩余度 = 1 - $H(X)/logr$

信源剩余度

- **关于剩余度的讨论**

提高无损信道信息传输率的研究就等于减少信源剩余度的研究

$\{S: S_1, S_2 \cdots S_R\}$ $\{X: X_1, X_2 \cdots X_r\}$



信源冗余度 $= 1 - \dfrac{H_\infty}{logR}$ $\Rightarrow$ 信源编码后 $\Rightarrow$ $\begin{cases} H(X) \to \log r \\ P(X)\text{等概分布} \end{cases}$

信道冗余度 $= 1 - \dfrac{H(X)}{logr}$ $\Rightarrow$ 信道冗余度 $= 1 - \dfrac{logr}{logr} = 0$