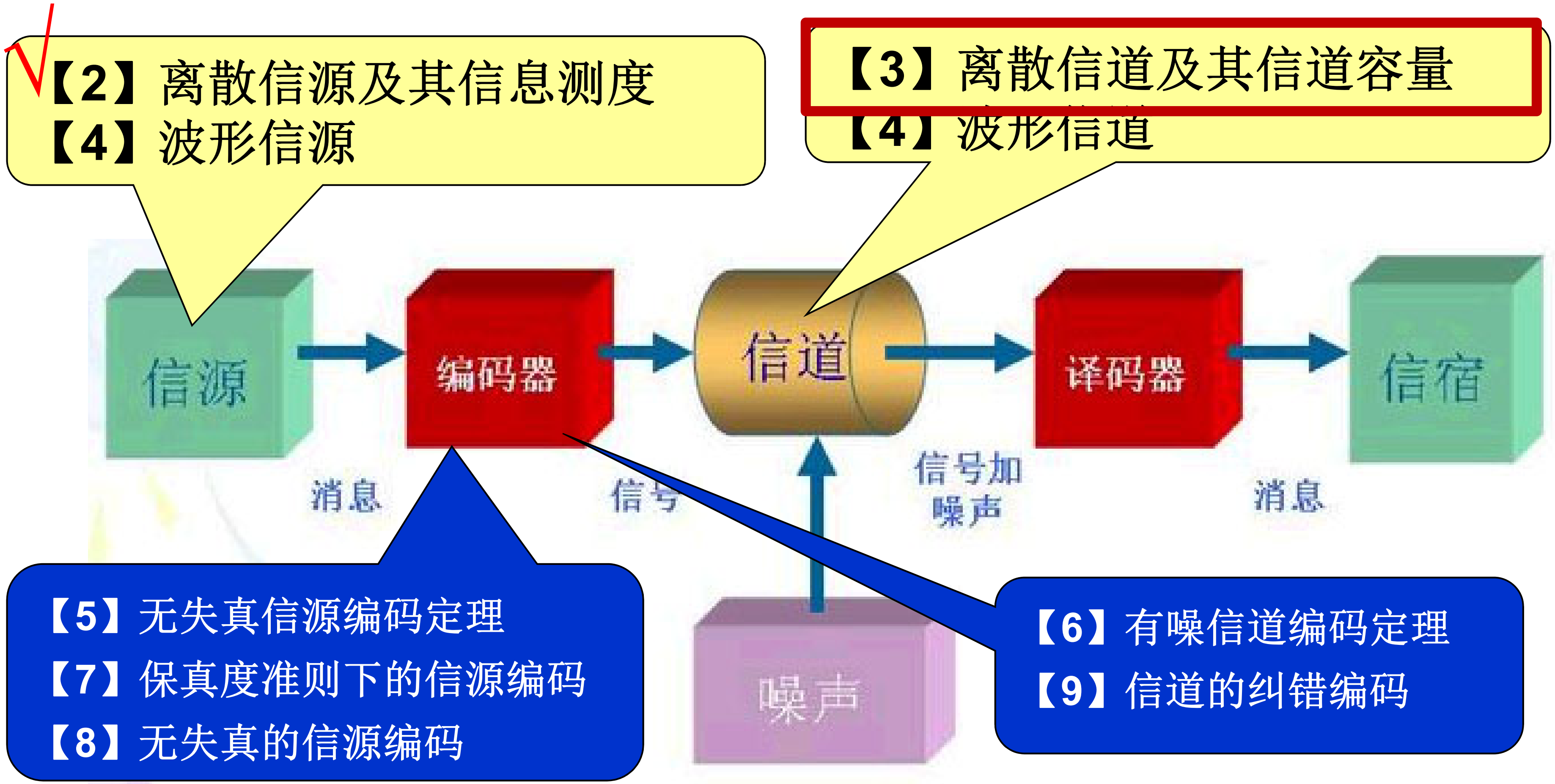
第三章 离散信道及其信道容量

2022/5/9

信息论的旅程

- 从本章开始，我们将从有效而可靠地传输信息的观点出发，对组成信息传输系统的各个部分进行讨论。



本章的研究内容★

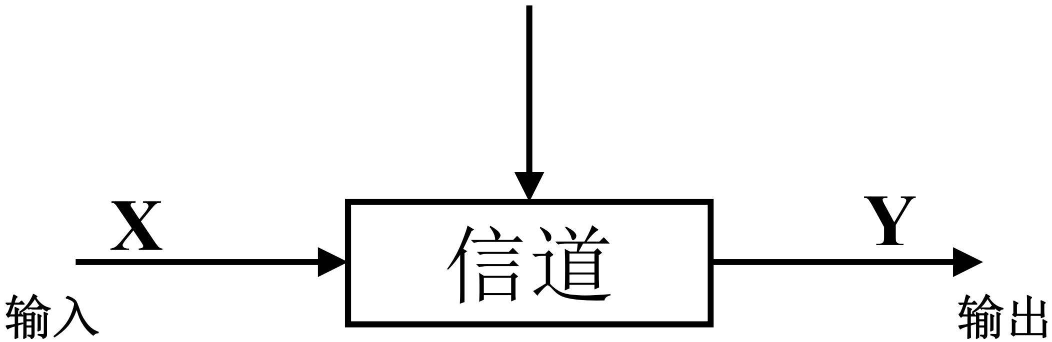
- 什么是信道?
  - 信号的传输媒介，是传送信息的物理通道。
- 信道的主要研究内容
  - 信道建模 – 信道的统计特性描述
  - 信道传输信息的能力 – 信道容量
  - 在有噪信道中能不能实现可靠传输?

离散信道及其容量 – 主要内容

- 信道的数学模型及其分类
  - 数学模型
  - 信道的分类
- 互信息及其性质
- 离散单符号无记忆信道
- 离散无记忆扩展信道
- 信道的组合 – 级联信道
- 信道容量

1.1、离散信道的数学模型

- 假定：传输特性已知



$$P(Y|X)$$

$$\{\mathbf{X},\ \ P(Y|X),\ \ \mathbf{Y}\}$$

1.2、信道分类

- 按输入/输出信号的幅度和时间特性划分：

幅度	时间	信道分类名称
离散	离散	离散信道（数字信道） 信息论中讨论纠错编码
连续	离散	连续信道
连续	连续	模拟信道（波形信道） 实际信道、有重要意义
离散	连续	（理论和实用价值均很小）

1.2、信道分类（续）

- 按输入/输出之间的记忆性来划分：
  - 无记忆信道：信道在某时刻的输出只与信道该时刻的输入有关，而与信道其他时刻的输入、输出无关。
  - 有记忆信道：信道在某时刻的输出与其他时刻的输入、输出有关。
- 根据信道的输入/输出是否是确定关系可分为:
  - 有干扰信道
  - 无干扰信道

1.2、信道分类（续）

- 信道的统计特性是否随时间改变而改变：
  - 平稳信道（恒参信道、时不变信道，如卫星通信）
  - 非平稳信道（随参信道、时变信道，如短波通信）
- 根据输入/输出的个数可分为：
  - 两端信道（单用户信道）：一个输入一个输出的单向通信。
  - 多用户信道：双向通信或三个或更多个用户之间相互通信的情况，例如多元接入信道、广播信道、网络通信信道等。
- 本章着重讨论单用户信道，以无记忆、恒参的有干扰、离散信道为重点。

离散信道及其容量 – 主要内容

- 信道的数学模型
- 互信息及其性质
- 离散单符号无记忆
- 离散无记忆扩展信道
- 信道的组合 – 级联信道
- 信道容量

1. 互信息定义与性质
2. 条件互信息
3. 平均互信息定义与性质

2.1、互信息 – 预备知识

信源 → 信道 → 信宿

噪声 → 信道

通信系统的简化模型

互信息的引入？

- 信源发出的消息，经过有噪信道传递到信宿。
  - 设*X*为信源发出的离散消息集合；*Y*为信宿收到的离散消息集合。

- 相关的概率知识
  - 先验概率：信源发出消息 x_i 的概率 $p(x_i)$
  - 后验概率：信宿收到消息 y_j 后推测信源发出 x_i 的概率，即条件概率 $p(x_i \mid y_j)$

2.1、互信息 － 定义

- 对离散随机事件集X, Y，互信息 $I(x_i;y_j)$为事件y_j所给出关于事件x_i的信息，

即 ★ $$I(x_i;y_j)=I(x_i)-I(x_i\mid y_j)=\log\frac{p(x_i\mid y_j)}{p(x_i)}$$

- 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

2.1、互信息 － 例题3.1

- 设 e 表示“降雨”，f 表示“空中有乌云”，且 p(e)=0.125，$p(e|f)=0.8$，求解：
  1. “降雨”的自信息 $I(e)$
  2. “空中有乌云”条件下“降雨”的自信息 $I(e|f)$
  3. “无雨”的自信息 $I(\bar{e})$
  4. “空中有乌云”条件下“无雨”的自信息 $I(\bar{e}|f)$
  5. “降雨”与“空中有乌云”的互信息 $I(e;f)$
  6. “无雨”与“空中有乌云”的互信息 $I(\bar{e};f)$

- 提示：通过概率空间完备性来完善概率求解！
- 3、0.322、0.193、2.322、2.678、-2.129

2.1、互信息 – 重要性质

- "互信息"的引出，使信息传输进入了定量分析，是信息论发展的一个重要的里程碑。

- 互信息的主要性质 ★
  – 互信息的互易性；
  – 互信息可为零；
  – 互信息可正可负；
  – 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；

2.1、互信息 – 互信息的性质

- 互信息的互易性 – 对称性
  - 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。

- 互信息可为零：表明 x_i 和 y_j 之间不存在统计约束关系，从 y_j 得不到关于 x_i 的任何信息，反之亦然。

2.1、互信息 – 互信息的性质（续一）

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

– 先验概率：$p(x_i)$　　后验概率：$p(x_i\mid y_j)$

- 互信息可正可负可为零
  - 何时为正？何时等于发送端信息量？
  - 何时为零？
  - 何时为负？

2.1、互信息 － 互信息的性质（续二）

- 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

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

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

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。

2.2、互信息 – 条件互信息

- 定义：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：

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

➢ 条件互信息和互信息的区别：
先验概率和后验概率都是在某一特定条件下的取值。

- 联合集*XYZ*中，x_i 与 $y_j z_k$ 的互信息量：

$$I(x_i; y_j z_k) = \log \frac{p(x_i|y_j z_k)}{p(x_i)} = \log\left[\frac{p(x_i|y_j z_k)}{p(x_i)} * \frac{p(x_i|y_j)}{p(x_i|y_j)}\right]$$

$$= I(x_i; y_j) + I(x_i; z_k|y_j)$$

2.3、平均互信息 － 定义

- 平均互信息：是互信息 $I(x_i;y_j)$ 在联合概率空间$P(XY)$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。
- 求解公式如下：

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

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

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

$I(X;y_j)$

2.3、平均互信息 － 定义（续）

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

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

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

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

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

条件熵：给定随机变量Y后，对随机变量X仍存在的不确定度

同样可以同三个角度来分析：输出端、输入端、通信系统总体

上堂课程主要内容复习

- 信道及其主要研究内容
- 信道的数学模型及分类
  - 重点讨论单用户信道，以无记忆、恒参、有干扰、离散信道为重点.
- 互信息的定义、性质
  - 条件互信息
- 平均互信息的定义

2.3、平均互信息 － 分析

观察者站在输出端

$$I(X;Y)=H(X)-H(X\mid Y)$$

- ***H*(*X*|*Y*)** —信道疑义度/损失熵/后验熵。 表示收到变量*Y*后，对随机变量*X*仍然存在的不确定度。代表了在信道中损失的信息。
- ***H*(*X*)** —*X*的先验不确定度/无条件熵/先验熵。
- ***I*(*X*;*Y*)**—收到*Y*前、后关于*X*的不确定度减少的量。从*Y*获得的关于*X*的平均信息量。

$$\boxed{\text{平均互信息}=\text{先验熵}-\text{后验熵}}$$

2.3、平均互信息 – 分析（续）

观察者站在输入端

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

- $H(Y|X)$—噪声熵。表示发出随机变量X后，对随机变量Y仍然存在的平均不确定度，可以理解为噪声引入的信息量（$Y=N+X$）。如果信道中不存在任何噪声，发送端和接收端必存在确定的对应关系，发出X后必能确定对应的Y，而现在不能完全确定对应的Y，这显然是由信道噪声所引起的。
- $I(Y;X)$ —收到Y后获得的信息量减去额外的信息量（噪声带来的）。

平均互信息=接收符号熵−噪声熵

2.3、平均互信息 － 分析（续）

观察者站在通信系统总体立场上

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

- 通信前把X和Y看成两个相互独立的随机变量，整个系统的先验不确定度为X和Y的联合熵$H(X)+H(Y)$；
- 通信后把信道两端出现X和Y看成是由信道的传递统计特性联系起来的，具有一定统计关联关系的两个随机变量，这时整个系统的后验不确定度由$H(XY)$描述。
- $I(X;Y)$ —通信前、后整个系统不确定度减少量，即经系统无失真传输（流通）的平均信息量。

2.3、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（左*X*右*Y*）

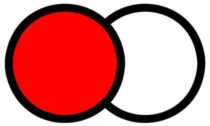
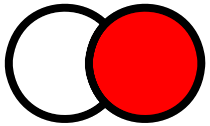
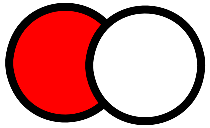
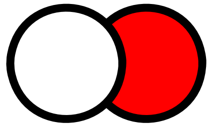
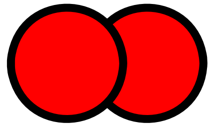
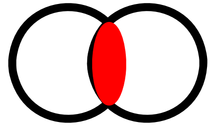
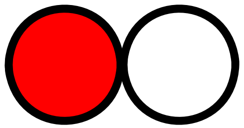
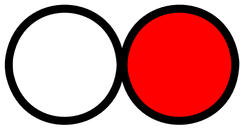
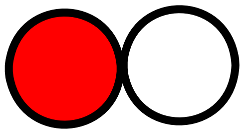
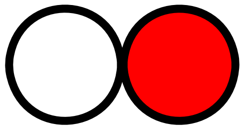
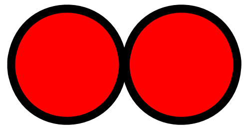
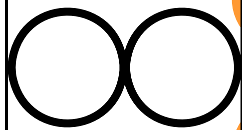
图示						
符号	$H(X)$	$H(Y)$	$H(X\|Y)$	$H(Y\|X)$	$H(XY)$	$I(X;Y)$

表2. 随机变量*X*、*Y*之间相互独立时

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

2.3 维拉图的练习 – 思考

- 几种特殊情况下的维拉图分析
  - 两个圈完全分离
  - 两个圈完全重合
  - 嵌入的情况：X在内
  - 嵌入的情况：Y在内

- 另外：若已知$H(X|YZ)=0$，可否画出对应的维拉图？

- 注意：维拉图不能用于证明类题目！
- 只可用于判断或分析

2.3、平均互信息的性质

1. 非负性★
   - 互信息不满足非负性
2. 互易性（对称性）
3. 凸函数性
4. 极值性

2.3、平均互信息的性质 – 非负性

$$\boxed{\ln x \le x-1 \qquad (x>0)}$$

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

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

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

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

2.3、平均互信息的性质 – 非负性的分析

- 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。

2.3、平均互信息的性质 – 对称性

证明

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

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

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

$$=I(Y;X)$$

2.3、平均互信息的性质 – 凸函数性

- 平均互信息量$I(X;Y)$是输入信源概率分布$p(x)$的上凸函数，是信道转移概率分布$p(y|x)$的下凸函数。

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

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

$$\Rightarrow \quad I(X;Y)=f[p(x),p(y|x)]$$

2.3、平均互信息的性质 – 极值性

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

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

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

- 说明：从输出端看，收到输出符号后平均每个符号获得的关于输入*X*的信息量不可能超过*X*自身的平均自信息。
- 最好情况、最差情况

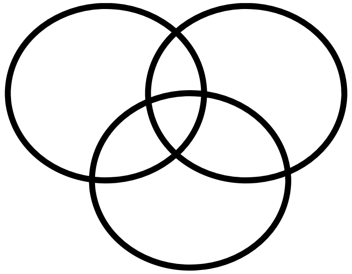
2.3、平均联合互信息

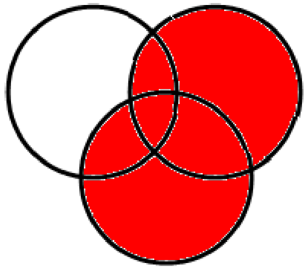
- 将联合互信息$I(x;yz)$在概率空间XYZ中求统计平均　　　　- 推导过程的区间变化

$$I(X;YZ)=E\left[I(x;yz)\right]=\sum_{X}\sum_{Y}\sum_{Z}p(xyz)\log\frac{p(x|yz)}{p(x)}$$

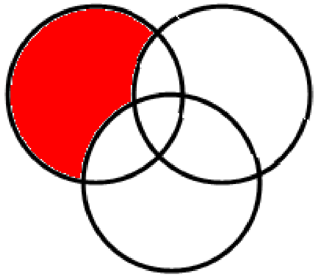
$$=H(X)-H(X|YZ)$$

-- 维拉图表示

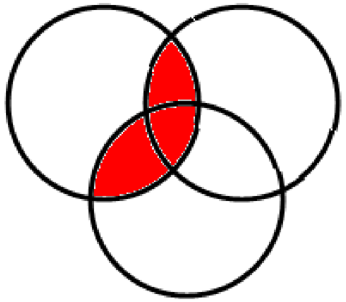




$H[YZ]$



$H[X/YZ]$



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

2.3、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 -- 维拉图表示

$$I(X;Z|Y)=E[I(x;z|y)]$$

$$=\sum_X\sum_Y\sum_Z p(xyz)\log\frac{p(x|yz)}{p(x|y)}=\sum_X\sum_Y\sum_Z p(xyz)\log\left(\frac{p(x|yz)}{p(x|y)}*\frac{p(x)}{p(x)}\right)$$

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

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

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

表示收到Y后再收到Z所多获得的关于X的信息量

2.3 维拉图的练习

$$H(XYZ)=H(XZ)+H(Y|X)-I(Z;Y|X)$$

- 用平均互信息与熵间关系证明
- 用定义公式证明

- 维拉图分析（不能用维拉图做证明题！）

2.3 平均互信息的扩展讨论

- 事物是普遍联系的，随机变量之间也存在相关关系

随机变量X → 某种操作或联系? → 随机变量Y

如何从信息的角度刻画X与Y之间的相关程度?

- 单独观察X，得到的信息量是$H(X)$
- 已知Y之后，X的信息量变为$H(X|Y)$
- 了解了Y之后，X的信息量减少了$H(X)-H(X|Y)$

这个减少量是得知Y取值之后提供的关于X的信息

信息论在数据挖掘中的应用-（决策树）

离散信道及其容量 – 主要内容

- 信道的数学模型及其分类
- 互信息及其性质
- 离散单符号无记忆信道
- 离散无记忆扩展信道
- 信道的组合
- 信道容量

1.单符号离散信道
2.信道疑义度
3.平均互信息
4.与其它熵的关系

3.1 离散无记忆信道的定义

- 离散无记忆信道（*DMC*），满足

$$p(\mathbf{Y}|\mathbf{X})=\prod_{i=1}^{N}p(Y_i|X_i)$$

- 离散无记忆恒参（平稳）信道，满足

$$p(Y_n=b_j|X_n=a_i)=p(Y_m=b_j|X_m=a_i)$$

- 注：若无特殊说明，所讨论离散无记忆信道均为平稳信道。因此，离散无记忆信道的研究只需研究单个符号的传输。

3.1、单符号离散信道- 数学模型

输入 **X** → 信道 → **Y** 输出

$\{X_1X_2\cdots X_N\}$ $\qquad$ $\{Y_1Y_2\cdots Y_N\}$

$X_i\in\{a_1,a_2,\cdots a_r\}$ $\qquad$ 噪声 $P(Y|X)$ $\qquad$ $Y_i\in\{b_1,b_2,\cdots b_s\}$

其中：

X 为一维随机变量，输入符号集 $A=\{a_1,a_2,\cdots a_r\}$

Y 为一维随机变量，输入符号集 $B=\{b_1,b_2,\cdots b_s\}$

信道传递/转移概率

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

$$(i=1,2,\cdots,r\quad j=1,2,\cdots,s)$$

$$\boxed{\sum_{j=1}^{s}p(b_j|a_i)=1}$$

3.1、单符号离散信道 – 信道矩阵

● 信道传递概率可以用信道矩阵**P**来表示：

$$\mathbf{P}=\begin{bmatrix} p(b_1|a_1) & p(b_2|a_1) & \cdots & p(b_s|a_1) \\ p(b_1|a_2) & p(b_2|a_2) & \cdots & p(b_s|a_2) \\ \vdots & \vdots & \ddots & \vdots \\ p(b_1|a_r) & p(b_2|a_r) & \cdots & p(b_s|a_r) \end{bmatrix}=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rs} \end{bmatrix}$$

描述信道特性

----> 信道中存在干扰的程度

$$\boxed{\sum_{j=1}^{s} p_{ij} = 1}$$

3.1、单符号离散信道 － 例题3.2

- 二元对称信道（*BSC*）输入符号集 $A=\{0,1\}$, 输出符号集 $B=\{0,1\}$，$r=s=2$。传递概率：

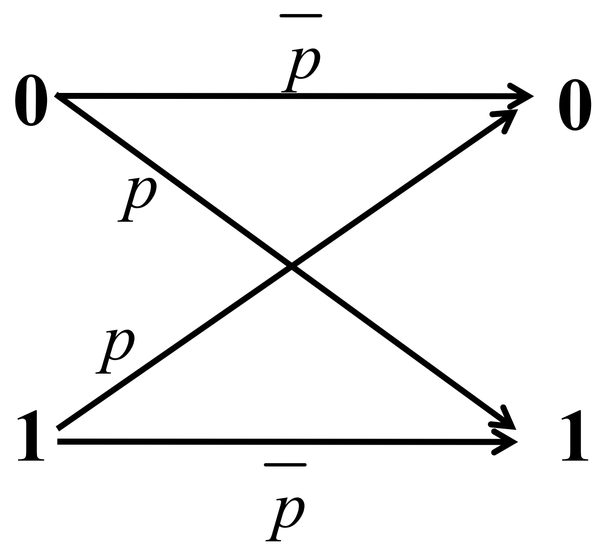
$$p(0|0)=\bar{p}$$

$$p(0|1)=p$$

$$p(1|0)=p$$

$$p(1|1)=\bar{p}$$

$$\mathbf{p}=\begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$



信道转移概率图

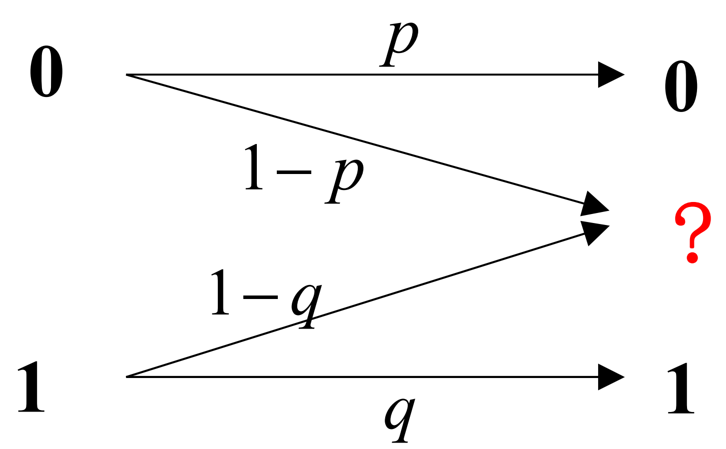
3.1、单符号离散信道 － 例题3.3

- 二元删除信道（*BEC*）。 输入符号集A={0,1}，符号输出集B={0,?,1}，$r = 2$，$s = 3$。

信道矩阵为：

$$\mathbf{p} = \begin{bmatrix} p & 1-p & 0 \\ 0 & 1-q & q \end{bmatrix}$$

$$\sum_{j=1}^{s} p\left(b_j \mid a_i\right) = 1$$



信道转移概率图

……是一种实际使用的信道……

3.1、单符号离散信道 － 例题3.3（续）

输入：以正负方波信号分别代表“0”和“1”

输出：由于信道中随机噪声干扰，输出端是受干扰后的方波信号，用$R(t)$表示。

用积分

$$I=\int R(t)dt$$

判断发送的信号是“0”还是“1”：

－若I是正值且大于某一限值，判“0”；

－若I是负值且小于某一限值，判“1”；

－若I绝对值小，不能判断，删除“?”；

前提：信道干扰不是很严重

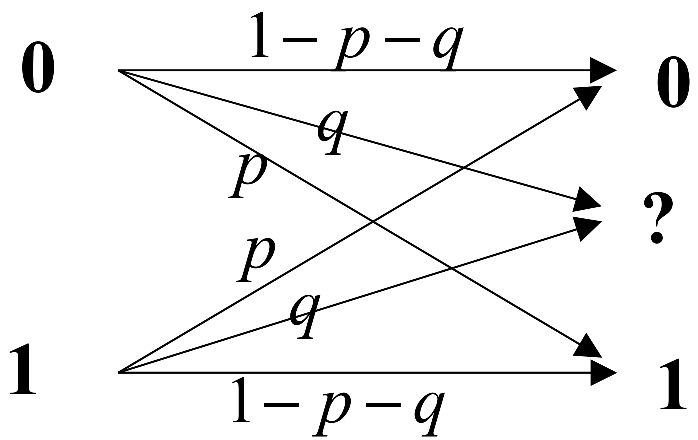
3.1、单符号离散信道 – 练习3.1

- 二元对称消失信道

  输入符号集A={0, 1}，

  符号输出集B={0, ?, 1}，

  写出对应的信道矩阵。



信道转移概率图

信道矩阵为：

$$\mathbf{p} = \begin{bmatrix} 1-p-q & q & p \\ p & q & 1-p-q \end{bmatrix}$$

3.1、离散无记忆信道 － 常用概率公式

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

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、离散无记忆信道 － 常用概率公式（续）

$$p\left(a_i \mid b_j\right)=\frac{p\left(a_i b_j\right)}{p\left(b_j\right)}=\frac{p\left(a_i\right)p\left(b_j \mid a_i\right)}{\sum_{i=1}^{r} p\left(a_i b_j\right)}=\frac{p\left(a_i\right)p\left(b_j \mid a_i\right)}{\sum_{i=1}^{r} p\left(a_i\right)p\left(b_j \mid a_i\right)}$$

5) 输出符号概率

$$p\left(b_j\right)=p\left(Y=b_j\right) \quad j=1,\cdots,s$$

用矩阵形式表示

$$\left[p\left(b_1\right) \quad p\left(b_2\right) \quad \cdots \quad p\left(b_s\right)\right]=\left[p\left(a_1\right) \quad p\left(a_2\right) \quad \cdots \quad p\left(a_r\right)\right]\times P_{Y|X}$$

$$P_Y = P_X \times P_{Y|X}$$

3.1、离散无记忆信道 － 常用概率公式（续）

- 结论
  - 有关信道的概率都可由先验概率 $p(a_i)$ 和前向概率 $p(b_j|a_i)$ 表示。

$$
\left.\begin{array}{l} p(a_i) \\ p(b_j|a_i) \end{array}\right\} \Rightarrow \left\{\begin{array}{l} p(b_j)=\sum_{i=1}^{r} p(a_i b_j)=\sum_{i=1}^{r} p(a_i)\cdot p(b_j|a_i) \\ p(a_i|b_j)=\dfrac{p(a_i b_j)}{p(b_j)}=\dfrac{p(a_i)p(b_j|a_i)}{p(b_j)} \end{array}\right.
$$

3.2、信道的疑义度（条件熵）

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

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

- 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。

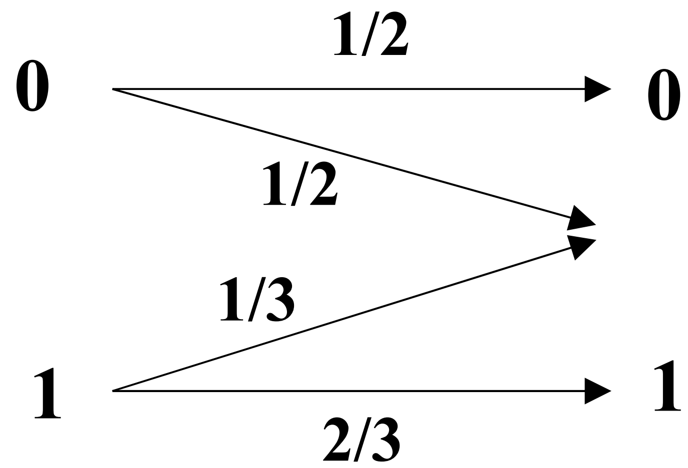
3.2、信道的疑义度 – 例题3.4

已知：输入集X的概率分布如下，求信道疑义度。

$$P_X = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

$$P_{Y|X} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

$$P_Y = P_X \times P_{Y|X} = \begin{bmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{bmatrix}$$



二元删除信道

3.2、信道的疑义度 – 例题3.4(续)

根据联合概率，可求后向（后验）概率

$$P_{X|Y}(0|0)=1 \qquad P_{X|Y}(1|0)=0$$
$$P_{X|Y}(0|?)=1/2 \qquad P_{X|Y}(0|?)=1/2$$
$$P_{X|Y}(0|1)=0 \qquad P_{X|Y}(0|1)=1$$

$$P_{X|Y}=\begin{bmatrix}1 & 0\\ 1/3 & 2/3\\ 0 & 1\end{bmatrix}$$

信道疑义度求解：

$$H(X|Y)=\sum_{XY}p_{xy}\log\frac{1}{p_{x|y}}=0.344 \quad bit/sym$$

$$H(X)=H\left(\frac{1}{4},\frac{3}{4}\right)=0.811 \quad bit/sym$$

3.3、平均互信息

- 平均互信息：$I(X;Y)=H(X)-H(X|Y)$

平均互信息表示接收到**Y**以后，平均每个符号所获得的关于输入变量**X**的信息量，是信道实际传输的信息量。

- 信源熵是信源输出的信息量，而真正被接收者收到的、关于信源内容的信息量则是平均互信息。

- 平均互信息的三种表达方式

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

- 平均互信息的性质：非负、互易、极值、凸函数性

3.3、平均互信息 – 凸函数性的定理

- 平均互信息是信源概率分布和信道传递概率的凸函数。$\boxed{I(X;Y)=f\left[P(X),P(Y|X)\right]}$
- 两个重要定理
  - 对于固定信道，即$P(Y|X)$不变的前提下，$I(X;Y)$是输入信源概率分布$P(X)$的“∩”型凸函数（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$P(Y|X)$的“∪”型凸函数（下凸函数）。

3.3、平均互信息 － 定理3.1

- 对于固定的信道，平均互信息量$I(X;Y)$是信源概率分布$P(X)$的上凸函数。

<u>证明思路</u>:

固定信道 $\longrightarrow$ $P(Y|X)$

上凸函数 $\longrightarrow$ 对应于“上凸”的证明表达

$P_1(X)$ $P_2(X)$ $I[P_1(X)]$和$I[P_2(X)]$

$P(X)$ $P(X)=\theta P_1(X)+\bar{\theta}P_2(X)$ $I[P(X)]$

需证明: $\boxed{\theta I[P_1(X)]+\bar{\theta}I[P_2(X)]-I[P(X)]\le 0}$

$$\boxed{\theta I[P_1(X)] + \bar{\theta} I[P_2(X)] - I[P(X)] \le 0}$$

代入平均互信息公式:

$$I[P_1(X)] = \sum_{i=1}^{r}\sum_{j=1}^{s} p_1(x_i y_j)\log\frac{p(y_j \mid x_i)}{p_1(y_j)} \cdots\cdots(1)$$

$$I[P_2(X)] = \sum_{i=1}^{r}\sum_{j=1}^{s} p_2(x_i y_j)\log\frac{p(y_j \mid x_i)}{p_2(y_j)} \cdots\cdots(2)$$

$$I[P(X)] = \sum_{i=1}^{r}\sum_{j=1}^{s} \underline{p(x_i y_j)}\log\frac{p(y_j \mid x_i)}{p(y_j)} \cdots\cdots(3)$$

$$p(x_i y_j) = p(x_i)p(y_j \mid x_i) = [\theta p_1(x_i) + \bar{\theta} p_2(x_i)]p(y_j \mid x_i)$$

$$= \theta p_1(x_i)p(y_j \mid x_i) + \bar{\theta} p_2(x_i)p(y_j \mid x_i)$$

$$= \theta p_1(x_i y_j) + \bar{\theta} p_2(x_i y_j)$$

$$\boxed{\theta I[P_1(X)]+\bar{\theta} I[P_2(X)]-I[P(X)]\le 0}$$

$$\begin{aligned}
&\theta I[P_1(X)]+\bar{\theta} I[P_2(X)]-I[P(X)]\\
&=\theta\sum_{i=1}^{r}\sum_{j=1}^{s}p_1(x_iy_j)\log\frac{p(y_j\mid x_i)}{p_1(y_j)}+\bar{\theta}\sum_{i=1}^{r}\sum_{j=1}^{s}p_2(x_iy_j)\log\frac{p(y_j\mid x_i)}{p_2(y_j)}\\
&-\sum_{i=1}^{r}\sum_{j=1}^{s}p(x_iy_j)\log\frac{p(y_j\mid x_i)}{p(y_j)}\\
&=\underline{\theta\sum_{i=1}^{r}\sum_{j=1}^{s}p_1(x_iy_j)\log\frac{p(y_j)}{p_1(y_j)}}+\underline{\bar{\theta}\sum_{i=1}^{r}\sum_{j=1}^{s}p_2(x_iy_j)\log\frac{p(y_j)}{p_2(y_j)}}
\end{aligned}$$

分别讨论这两项

3.3、定理3.1证明（续）

- $f(x) = \log x$是“$\cap$”型凸函数

$$\theta \sum_{i=1}^{r} \sum_{j=1}^{s} p_1(x_i y_j) \log \frac{p(y_j)}{p_1(y_j)}$$

$$\le \theta \ \log \sum_{i=1}^{r} \sum_{j=1}^{s} p_1(x_i y_j) \frac{p(y_j)}{p_1(y_j)}$$

$$= \theta \ \log \sum_{j=1}^{s} \frac{p(y_j)}{p_1(y_j)} \sum_{i=1}^{r} p_1(x_i y_j)$$

$$= \theta \ \log \sum_{j=1}^{s} p(y_j) = 0$$

同样，可以用詹森不等式证明第二项也小于等于0

3.3、定理3.1证明（续）

$$I[\theta P_1(X)+\bar{\theta}P_2(X)]\geq\theta I_1[P_1(X)]+\bar{\theta}I_2[P_2(X)]$$

$\because\theta$ 和 $\overline{\theta}$ 均为0至1间的正数，且 $\theta+\overline{\theta}=1.$

$\therefore$ 根据上凸函数的定义：

自变量期望的函数值 $\geq$ 函数值的期望。

可得：

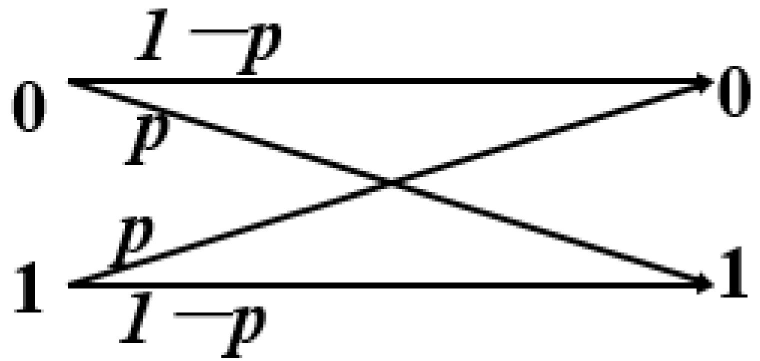
$I[P(X)]$是输入信源的概率分布$P(X)$的“$\cap$”型凸函数。

3.3、平均互信息 – 例题3.5

- 二进制对称信道，输入X的概率分布,即信源概率空间为：

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

信道如图：



$$\boxed{1 - p = \overline{p}}$$

试分析平均互信息与信源概率分布的关系

3.3、平均互信息 – 例题3.5（续）

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

$$=H(Y)-\sum_{i=1}^{2}\sum_{j=1}^{2}p(x_i)p(y_j|x_i)\log p(y_j|x_i)$$

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

$$=H(Y)-\underline{\sum_{i=1}^{2}p(x_i)}\{-[\bar{p}\log\bar{p}+p\log p]\}$$

$$=H(Y)-\{-[\bar{p}\log\bar{p}+p\log p]\}=H(Y)-H(p)$$

$$p(y_1=0)=p(x=0)p(0|0)+p(x=1)p(0|1)=\omega\bar{p}+\bar{\omega}p$$

$$p(y_2=1)=p(x=0)p(1|0)+p(x=1)p(1|1)=\omega p+\bar{\omega}\bar{p}$$

$$H(Y)=(\omega\bar{p}+\bar{\omega}p)\log\frac{1}{\omega\bar{p}+\bar{\omega}p}+(\omega p+\bar{\omega}\bar{p})\log\frac{1}{\omega p+\bar{\omega}\bar{p}}$$

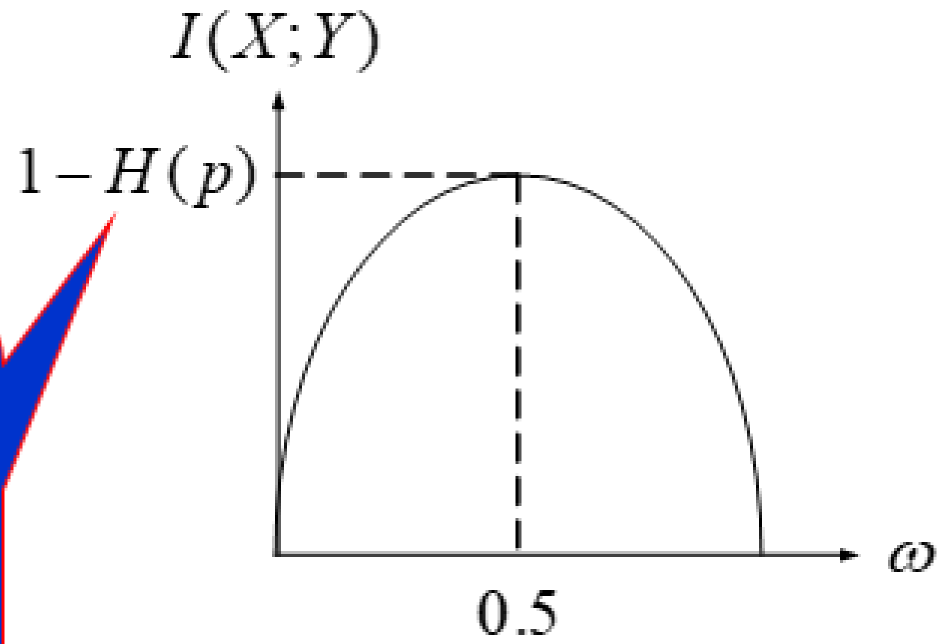
$$=H(\omega\bar{p}+\bar{\omega}p)$$

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

3.3、平均互信息 – 例题3.5（续）

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，$H(\omega\bar{p}+\bar{\omega}p)=H(\frac{1}{2})=1$ 平均互信息取得最大值。



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。

3.3、平均互信息 - 定理3.2

- 对于固定的信源分布，平均互信息$I(X;Y)$是信道传递概率$P(Y|X)$的下凸函数。

证明思路:

固定信源分布 → $P(X)$

下凸函数 → 对应于“下凸”的证明表达

$P_1(Y|X)$ $P_2(Y|X)$ $I[P_1(Y|X)]$和$I[P_2(Y|X)]$

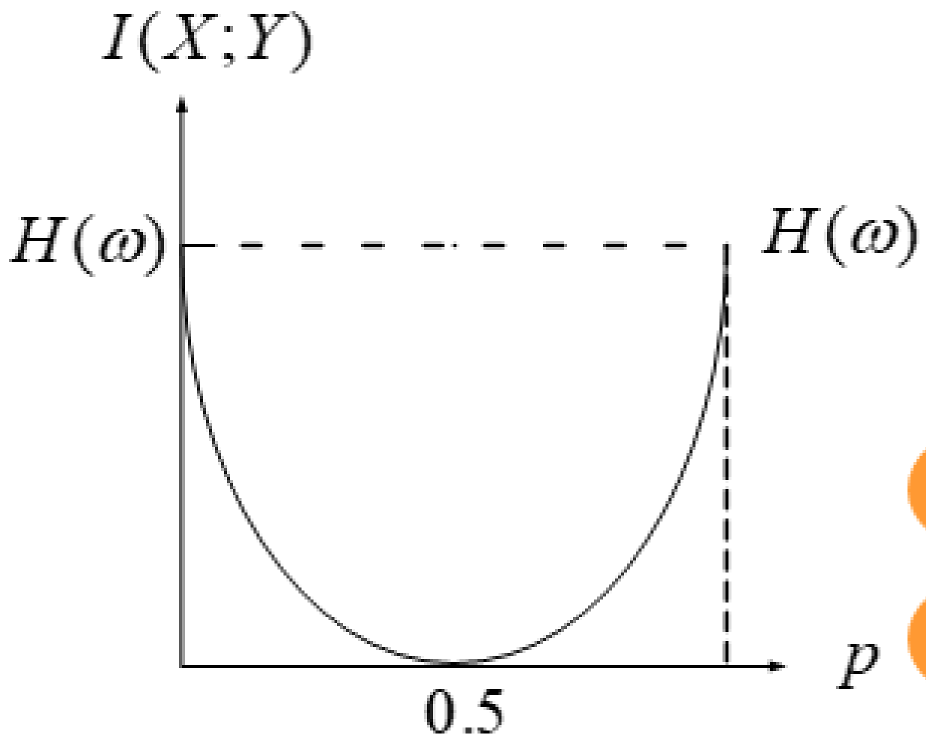
$P(Y|X)$ $P(Y|X)=\theta P_1(Y|X)+\bar{\theta}P_2(Y|X)$ $I[P(Y|X)]$

需证明: $\boxed{I[P(Y|X)]-\theta I[P_1(Y|X)]-\bar{\theta} I[P_2(Y|X)]\le 0}$

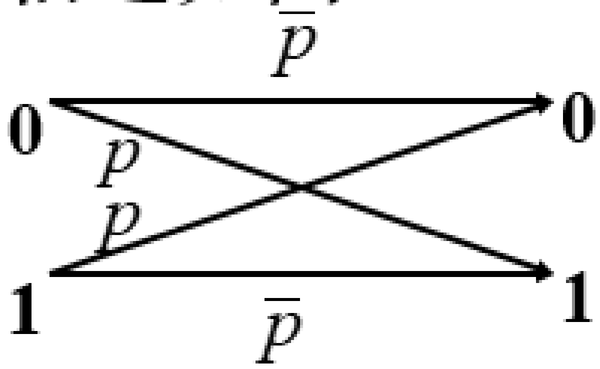
3.3、平均互信息 – 例题3.6

- 例（3.5续）：二进制对称信道，输入X概率分布，即信源概率空间为：

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



信道如图：



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

关于定理的说明

p＝0时，无噪信道：$I(X;Y)=H(\omega)$
p＝1时，无噪信道：$I(X;Y)=H(\omega)$

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。

物理意义：

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 – 例题3.7

- 掷骰子，如果结果是1、2、3、4，则抛一次硬币；如果结果是5或6，则抛两次硬币，试计算从抛硬币正面出现次数可以得出多少掷骰子所得点数的信息量。

根据题意，可定义：X，Y，$P(Y|X)$，求 $I(X;Y)$

掷骰子事件 --〉 X

结果1,2,3,4 --〉 $x_1 = 0$

结果5,6 --〉 $x_2 = 1$

$$\begin{bmatrix} X \\ P_X \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

抛硬币事件 --〉 Y

3.3、平均互信息 – 例题3.7（续）

信道矩阵

$$P_{Y|X}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{bmatrix}$$

***Y*集合元素**

0：0次正面

1：1次正面

2：2次正面

输出符号集Y的分布

$$P_Y=P_X\bullet P_{Y|X}=\begin{bmatrix}\frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4}\end{bmatrix}=\begin{bmatrix}\frac{5}{12} & \frac{1}{2} & \frac{1}{12}\end{bmatrix}$$

$$H(Y)=$$

$$H(Y|X)=$$

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

3.4、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

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

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

$$I(X;Y) = I(Y;X) \ge 0$$

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

- 平均互信息的凸函数性

上堂课程主要内容复习

- 单符号无记忆信道
  - 基本概念
    - 信道传递/转移概率、信道矩阵、疑义度、平均互信息
  - 平均互信息：真正被接收者收到的、关于信源内容的信息量。与熵的区别！
  - 常用概率 $p(x)$ $p(y|x)$ $p(xy)$ $p(x|y)$ $p(y)$
    - 信道输入概率和信道传递概率
  - 平均互信息的凸函数性 $\boxed{I(X;Y)=f[P(X),P(Y\mid X)]}$
    - 是信道输入分布和信道传递概率的凸函数

离散信道及其容量 － 主要内容

- 信道的数学模型及其分类
- 互信息及其性质
- 离散无记忆信道
- 离散信道中的序列传递
- 信道的组合
- 信道容量

1. （选）两个重要定理
2. 无记忆扩展

4.1、一般离散信道序列传递的平均互信息计算

- 若信道输入/出分别是N长序列$\mathbf{X}$和$\mathbf{Y}$，且信道无记忆，则存在 $I(\mathbf{X};\mathbf{Y}) \le \sum_{k=1}^{N} I(X_k;Y_k)$
  - 若信源为无记忆信源时，等号成立

分析：

无记忆信道；<=；分别列出表达式

$$p(\mathbf{y}\,|\,\mathbf{x})=\prod_{i=1}^{N} p(y_i\,|\,x_i) \qquad p(\beta_h\,|\,\alpha_k)=\prod_{i=1}^{N} p\left(b_{h_i}\,|\,a_{k_i}\right)$$

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

$$I(\mathbf{X};\mathbf{Y}) = \sum_{i=1}^{r^N}\sum_{j=1}^{s^N} p(\alpha_i\beta_j)\log\frac{p(\beta_j|\alpha_i)}{p(\beta_j)}$$

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

$$\sum_{k=1}^{N} I(X_k;Y_k) = \sum_{k=1}^{N}\left\{\sum_{i_k=1}^{r}\sum_{j_k=1}^{s} p(a_{i_k}b_{j_k})\log\frac{p(b_{j_k}\mid a_{i_k})}{p(b_{j_k})}\right\}$$

$$= \sum_{i_1=1}^{r}\sum_{j_1=1}^{s} p(a_{i_1}b_{j_1})\log\frac{p(b_{j_1}\mid a_{i_1})}{p(b_{j_1})} + \ldots + \sum_{i_N=1}^{r}\sum_{j_N=1}^{s} p(a_{i_N}b_{j_N})\log\frac{p(b_{j_N}\mid a_{i_N})}{p(b_{j_N})}$$

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

4.1、一般离散信道 - 序列传递的平均互信息计算 - 证明

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

$f(x)=\log x$是“∩”型凸函

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

$$\le \log\left\{\sum_{j_1=1}^{s}\cdots\sum_{j_N=1}^{s} p(b_{j_1}\ldots b_{j_N})\cdot\frac{p(b_{j_1})\ldots p(b_{j_N})}{p(b_{j_1}\ldots b_{j_N})}\right\}$$

$$=\log\left\{\sum_{j_1=1}^{s} p(b_{j_1})\cdots\sum_{j_N=1}^{s} p(b_{j_N})\right\}$$

信道无记忆 $\Rightarrow I(\mathbf{X};\mathbf{Y})-\sum_{k=1}^{N} I(X_k;Y_k)\le\log 1=0$

4.1、序列传递的平均互信息计算－ 说明

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

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

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

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

物理意义：

离散无记忆信道传递N长序列的平均互信息不会超过N个随机变量$X_1, X_2, \ldots, X_N$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

4.1、一般离散信道 - 序列传递的平均互信息计算

- 若信道的输入和输出分别是N长序列$\mathbf{X}$和$\mathbf{Y}$，且信源是无记忆的，则存在

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

当信道无记忆时，等号成立。

分析：无记忆信源；>=；分别列出表达式

$$p(\mathbf{x}) = \prod_{i=1}^{N} p(x_i) \qquad p(\alpha_k) = \prod_{i=1}^{N} p(a_{k_i}) \qquad (k = 1,2,\cdots,r^N)$$

4.1、序列传递的平均互信息计算 - 说明

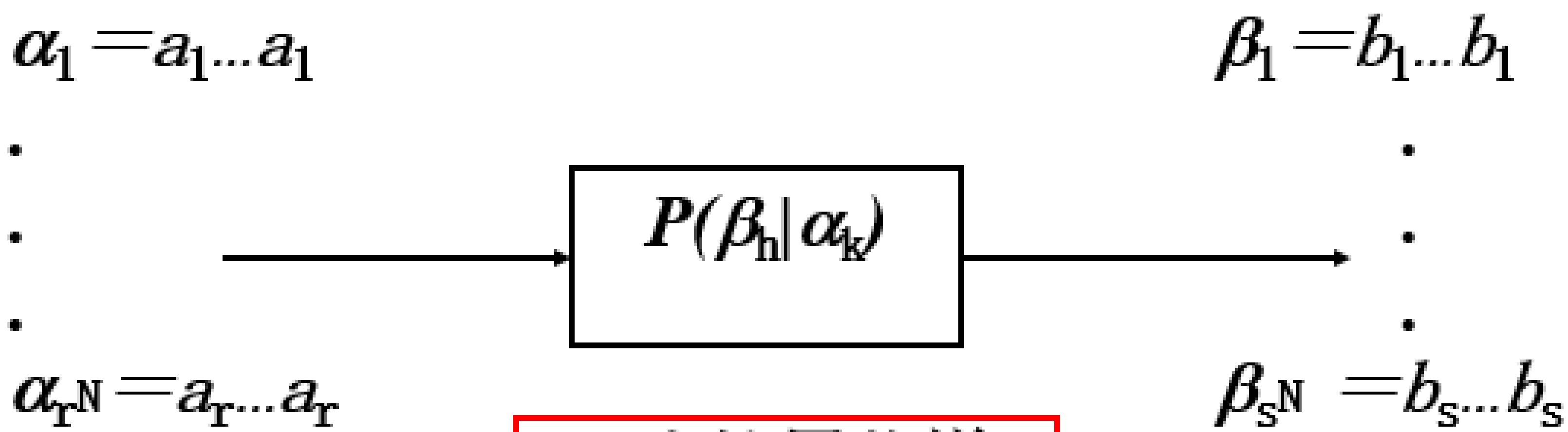
物理意义：

当信源无记忆时，平均互信息$I(X;Y)$不小于各变量单独通过单符号离散信道的平均互信息$I(X_i;Y_i)$之和。

当信道无记忆时两者相等，此时扩展信道$I(X;Y)$是未扩展信道平均互信息的N倍。

4.2、离散无记忆扩展信道 – N次扩展信道

- 一般离散信道输入和输出是一个**随机变量序列**
- **无记忆扩展**：每一个随机变量均取值于**同一**输入或输出符号集。



N次扩展信道

4.2、N次无记忆扩展信道 - 信道矩阵

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

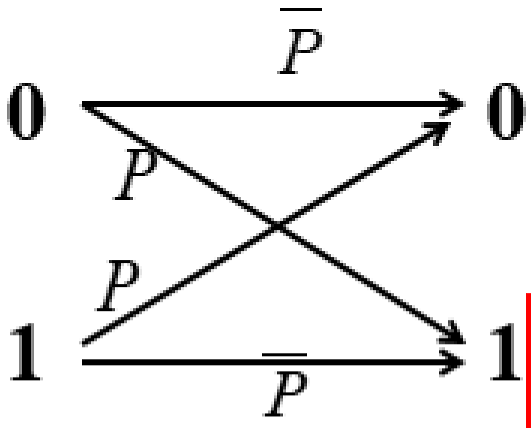
$$\pi_{ij} = p(\beta_j | \alpha_i) \qquad p_{ij} \geq 0; \quad \sum_{j=1}^{s^N} p_{ij} = 1 \qquad (i = 1, 2, \cdots, r^N)$$

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：

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

4.2、N次无记忆扩展信道 - 例题3.8

- 二元无记忆对称信道的二次扩展信道
  - 输入输出随机变量X和Y取值同一集合$A=\{0，1\}$



二次扩展的输入/出符号集合：

$$A^2 = B^2 = (00 \quad 01 \quad 10 \quad 11)$$

二次扩展信道的信道矩阵

$$
\begin{array}{c|cccc}
 & \underset{00}{\beta_1} & \underset{01}{\beta_2} & \underset{10}{\beta_3} & \underset{11}{\beta_4} \\
\hline
\alpha_1 = 00 & \bar{p}^2 & \bar{p}p & p\bar{p} & p^2 \\
\alpha_2 = 01 & \bar{p}p & \bar{p}^2 & p^2 & p\bar{p} \\
\alpha_3 = 10 & p\bar{p} & p^2 & \bar{p}^2 & \bar{p}p \\
\alpha_4 = 11 & p^2 & p\bar{p} & \bar{p}p & \bar{p}^2
\end{array}
$$

4.2 N次无记忆扩展信道 – 例题3.9

- 4个等概消息，编成的码字M_1=000，M_2=011，M_3=101，M_4=110，通过二元对称无记忆信道（p<0.5）传输。
- （1）事件“接收到第一个数字为‘0’”与发送M_1的互信息？
- （2）当“接收到第二个数字也是‘0’”时，求关于M_1的附加信息？
- （3）当“接收到第三个数字还是‘0’”时，又增加了多少关于M_1的消息？

例题3.9的推导思路

- 翻译题目
- 互信息
- 条件互信息

4.2、N次无记忆扩展信道 － 平均互信息

◆ N次扩展信道的平均互信息定义：

$$I(\mathbf{X};\mathbf{Y})=I\left(X^N;Y^N\right)$$

$$=H\left(X^N\right)-H\left(X^N\mid Y^N\right)=\sum_{X^NY^N}p\left(\alpha_k\beta_h\right)\log\frac{p\left(\alpha_k\mid\beta_h\right)}{p\left(\alpha_k\right)}$$

$$=H\left(Y^N\right)-H\left(Y^N\mid X^N\right)=\sum_{X^NY^N}p\left(\alpha_k\beta_h\right)\log\frac{p\left(\beta_h\mid\alpha_k\right)}{p\left(\beta_h\right)}$$

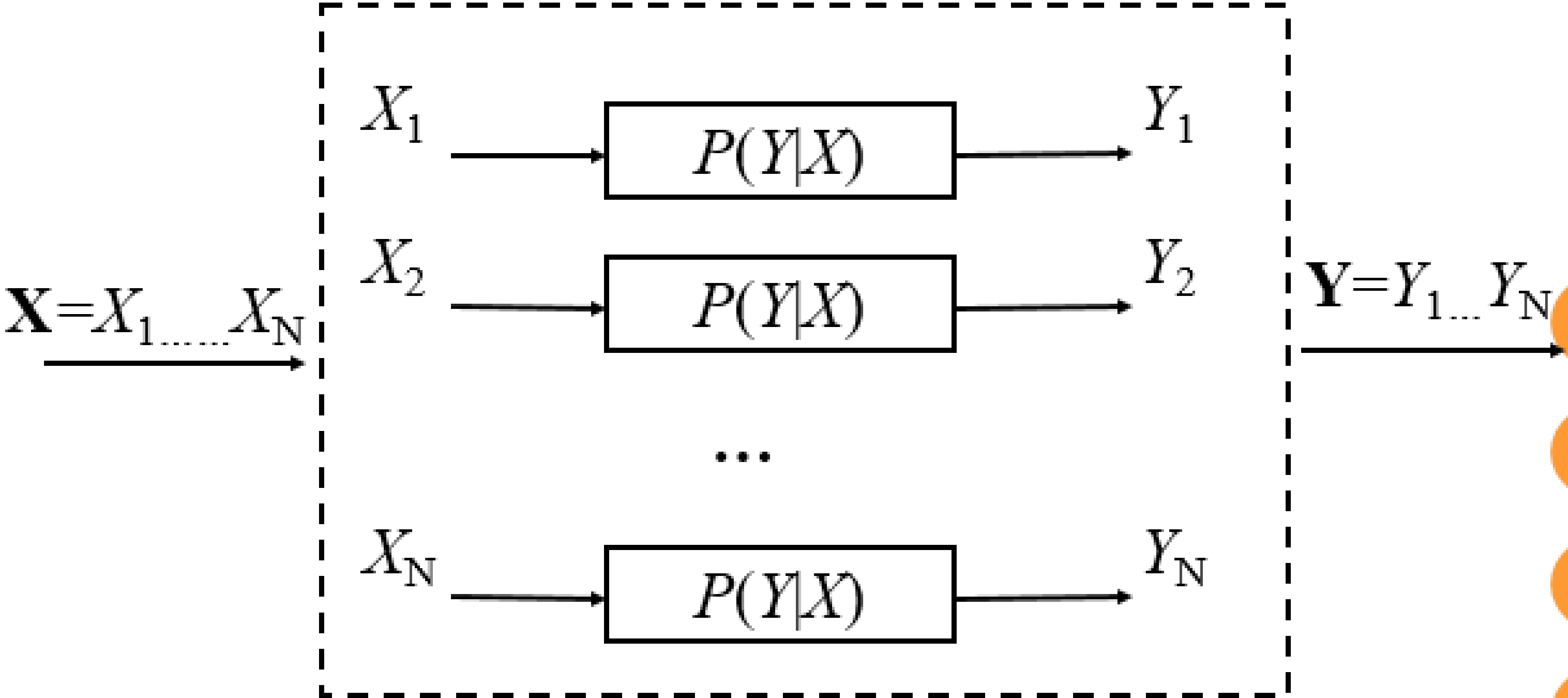
$$\left(k=1,2,\cdots\cdots,r^N\right),\qquad\left(h=1,2,\cdots\cdots,s^N\right)$$

信源无记忆时，N次无记忆扩展信道的平均互信息等于单个符号的平均互信息之和（单符号平均互信息的N倍）。

4.2、离散无记忆扩展信道 －讨论

- 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

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



4.2、离散无记忆扩展信道 - 讨论（续）

- 对于N次扩展信道，若
  - 信源无记忆，且输入序列中的每个分量都来自同一信源符号集，并具有相同分布
  - 经过相同的无记忆信道

则：信道输出随机序列中各符号之间亦无记忆，且下式成立

$$I(\mathbf{X};\mathbf{Y}) = N * I(X;Y)$$

– 对于无扰的一一对应的无记忆信道

$$I(X_i;Y_i) = H(X_i) \qquad H(\mathbf{X}) \le \sum_{i=1}^{N} H(X_i)$$

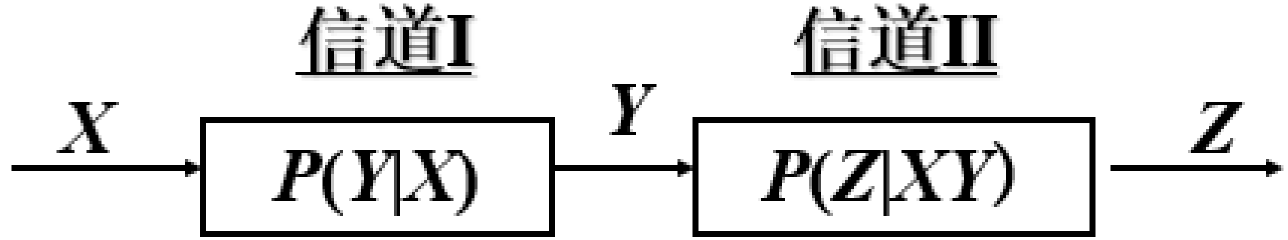
离散信道及其容量 – 主要内容

- 信道的数学模型及其分类
- 互信息及其性质
- 离散无记忆信道
- 离散无记忆扩
- 信道的组合
- 信道容量

1. 基本概念
2. 数据处理定理
3. 串联信道的信道矩阵

5.1、信道的组合

- 常见信道组合
  - 积信道：多个信道并行传送
  - **级联信道**：多个信道串行传送，*e.g.*无线电中继信道、通信系统的物理模型，……
  - 和信道：两个以上信道联合起来



$$A=\{a_1,a_2,\cdots,a_r\} \qquad B=\{b_1,b_2,\cdots,b_s\} \qquad C=\{c_1,c_2,\cdots,c_l\}$$

$$p(b_j \mid a_i) \qquad p(c_k \mid a_i b_j)$$

5.1、信道的组合 – 级联信道 – 平均互信息

I(*XY*;*Z*)：表示联合随机变量*XY*与随机变量*Z*之间的平均互信息。即，收到*Z*后，从*Z*中获得关于联合随机变量*XY*的平均信息量。

$$I(XY;Z)=\sum_{XYZ}p(xyz)\log\frac{p(z|xy)}{p(z)}$$

定理3.4 ★

$$I(XY;Z)\geq I(Y;Z) \qquad (1)$$

$$I(XY;Z)\geq I(X;Z) \qquad (2)$$

当且仅当

$p(z\mid xy)=p(z\mid y)$

$p(z\mid xy)=p(z\mid x)$时,(1),(2)式中等号分别成立

5.1、信道的组合 － 定理证明

$$\boxed{I(XY;Z) \geq I(Y;Z)}$$

$$I(Y;Z) - I(XY;Z)$$

$$= \sum_{YZ} p(yz)\log\frac{p(z|y)}{p(z)} - \sum_{XYZ} p(xyz)\log\frac{p(z|xy)}{p(z)}$$

$$= \sum_{XYZ} p(xyz)\log\frac{p(z|y)}{p(z)} - \sum_{XYZ} p(xyz)\log\frac{p(z|xy)}{p(z)}$$

$$= \sum_{XYZ} p(xyz)\log\frac{p(z|y)}{p(z|xy)}$$

当 $p(z|xy) = p(z|y)$ 的时候，等号成立

詹森不等式

$$\leq \log\sum_{XYZ} p(xyz)\frac{p(z|y)}{p(z|xy)}$$

$$= \log\sum_{XYZ} p(xy)p(z|y) = \log\sum_{YZ} p(y)p(z|y)$$

$$= \log\sum_{YZ} p(yz) = 0$$

5.1、信道的组合 – 定理证明

$$\boxed{I(XY;Z) \geq I(Y;Z)}$$

- 该定理的物理含义：
  - $P(Z|XY)=P(Z|Y)$：当随机变量Y确定后，随机变量Z值即可确定，而并不需要了解随机变量X。即，随机变量序列（XYZ）构成马尔可夫链。
  - 从Z中获取关于Y的平均互信息$I(Y;Z)$，一般不会超过从Z中获取关于联合随机变量XY的平均互信息$I(XY;Z)$，仅当随机变量序列（XYZ）是马氏链时，才相等。

* 一般的串联信道，随机变量序列（XYZ）可构成马氏链，即，Z与X没有直接的依赖关系。

5.2、信道的组合 – 定理3.5

- 若随机变量*X*, *Y*, *Z*构成一个马尔可夫链，则有

★
$$I(X;Z) \le I(Y;Z) \qquad (1)$$

$$I(X;Z) \le I(X;Y) \qquad (2)$$

证明：马尔可夫链 $\longrightarrow$ $p(z \mid xy) = p(z \mid y)$

$$\Downarrow$$

$$I(XY;Z) = I(Y;Z)$$

$$I(XY;Z) \ge I(X;Z)$$

$$I(Y;Z) \ge I(X;Z)$$

5.2、信道的组合 － 定理3.5（续）

$$\boxed{I(X;Z)\le I(X;Y)}$$

马尔可夫过程的逆，也满足马尔可夫性质

$$p(x\mid yz)=p(x\mid y)$$

同理可证：$I(X;Y)\ge I(X;Z)$

当$p(x\mid yz)=p(x\mid y)=p(x\mid z)$时，等号成立

说明：

1. 数据处理定理：信息不增性原理★
2. 级联信道上传输的最终信息量必然小于等于子信道各自分别传输的信息量
3. 若**XYZ**满足马氏链，则 **[P]＝[P$_{\mathrm{I}}$][P$_{\mathrm{II}}$]**

数据处理定理－讨论

1. 对接收到的信号/数据进行处理能否增加信息量？

   任何无源处理过程总是丢失信息的！

2. ?? 既然是信息是不增的，为何还要处理？
   - 如：通信系统的基本模型
3. 如何能增加可获取的信息？

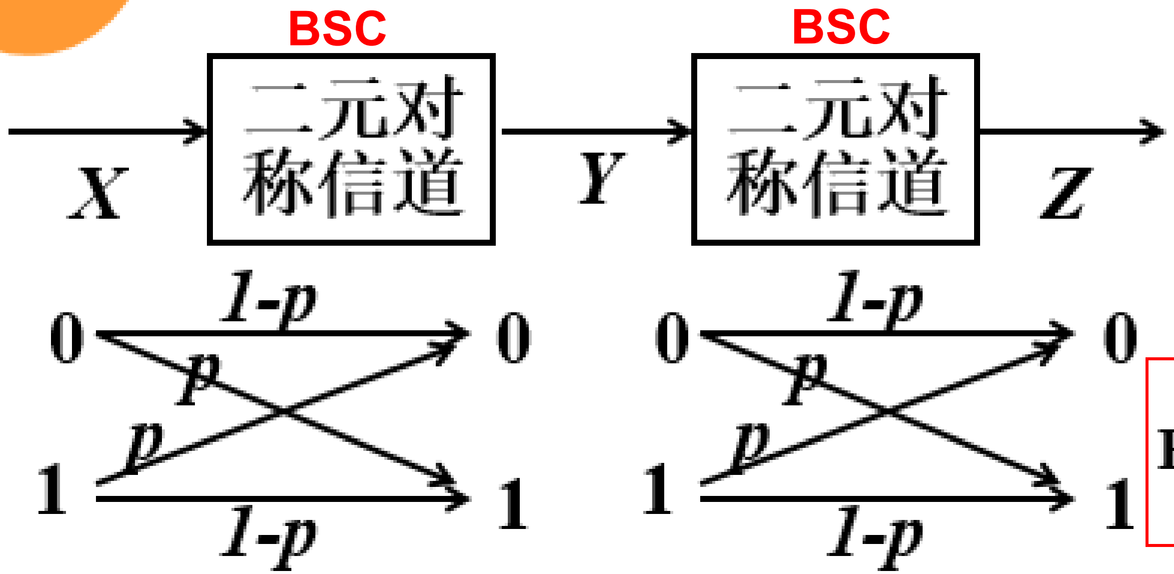
$$I(X;Y_1Y_2)=I(X;Y_1)+I(X;Y_2|Y_1)$$

数据处理定理（信息不增性原理） - 补充

$$I(X;Z) \le I(Y;Z) \qquad I(X;Z) \le I(X;Y)$$

- 如果将级联信道中的各个信道看成数据处理环节，则该定理含义为：经过各处理环节对数据进行处理后，一般只会增加信息的损失，最多保持原来所具有的信息，不可能比未经处理获得更多的信息。
- 在任何信息处理传输系统中，最后所获得的信息至多是信源所能提供的信息。一旦在某一过程丢失了一些信息，以后的系统无论如何处理，如不能触及到丢失信息过程的输入端，就不能再恢复已丢失的信息。

5.3、信道的组合 – 例题3.10两个二元对称信道的串联



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

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

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

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

5.3、信道的组合 – 例题3.10（续）

- 平均互信息

通过第一个信道 $I(X;Y)=1-H(p)$

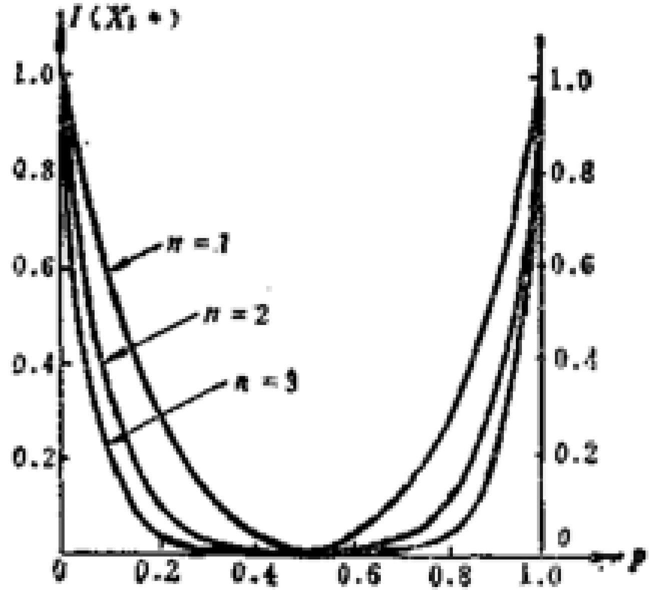
通过两个信道 $I(X;Z)=1-H[2p(1-p)]$

通过三个信道 $I(X;W)=1-H\left|3p(1-p)^2+p^3\right|$

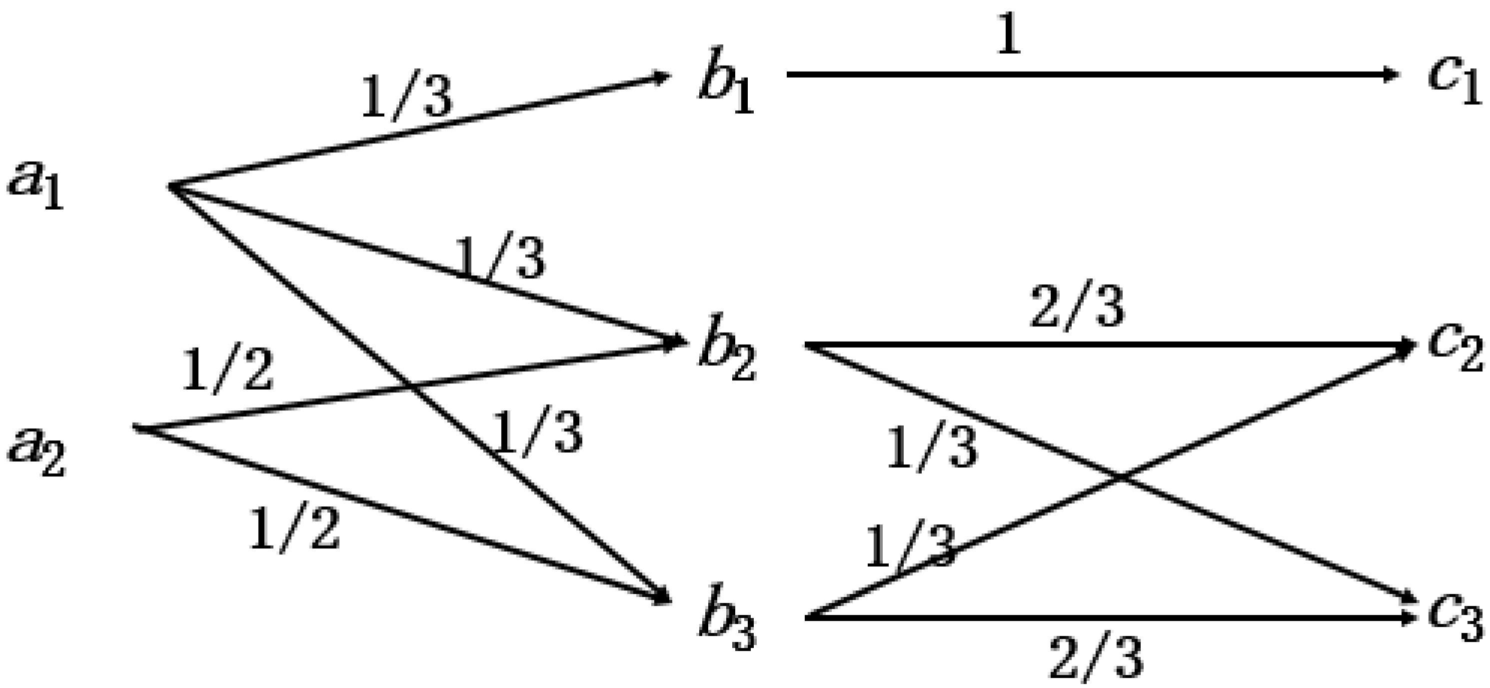
说明：

1. 信道串联后增加信息损失

2. 串联级数越多，损失越大



5.3、信道的组合 – 例题3.11



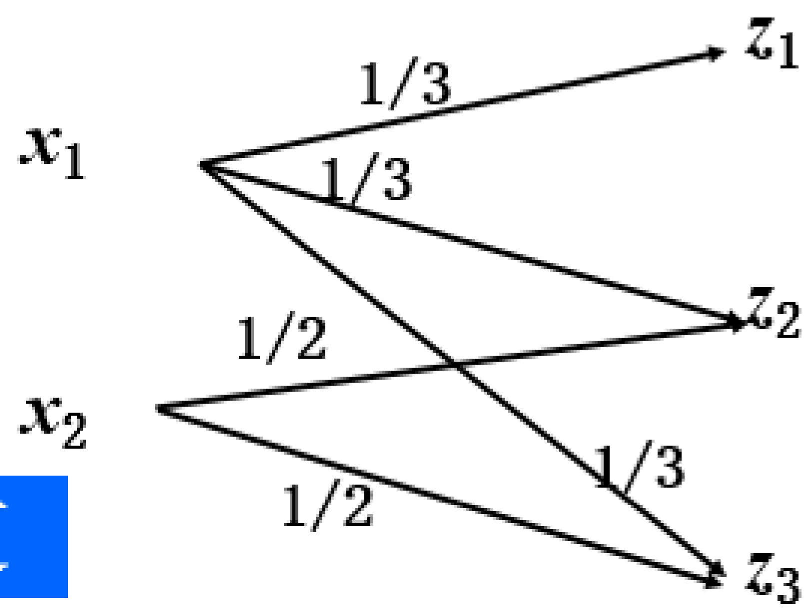
X、Y、Z满足马氏链的性质，求总的信道矩阵

5.3、信道的组合 – 例题3.11 （续）

$$\mathbf{P}_{Z|X} = \mathbf{P}_{Y|X} * \mathbf{P}_{Z|Y}$$

$$= \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

$$= \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$



可获得一个等价信道

离散信道及其容量 － 主要内容

- 信道的数学模型及其分类
- 互信息及其性质
- 离散无记忆
- 离散无记忆
- 信道的组合
- 信道容量

1. 信道容量C的定义
2. 特殊信道的C
3. 离散对称信道的C
4. 信道容量定理 及 准对称
5. 信源与信道匹配

6.1、信道容量 – 定义

1、信道容量的定义

- 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。

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

- 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：

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

6.1、信道容量 – 定义（续）

- 在信道确定的情况下，$I(X;Y)$ 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 $I(X;Y)$ 最大。定义这个<u>最大的信息传输率</u>为**信道容量**：

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

相应的输入概率分布被称为**最佳输入分布**。

6.1、信道容量 － 说明

- 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。

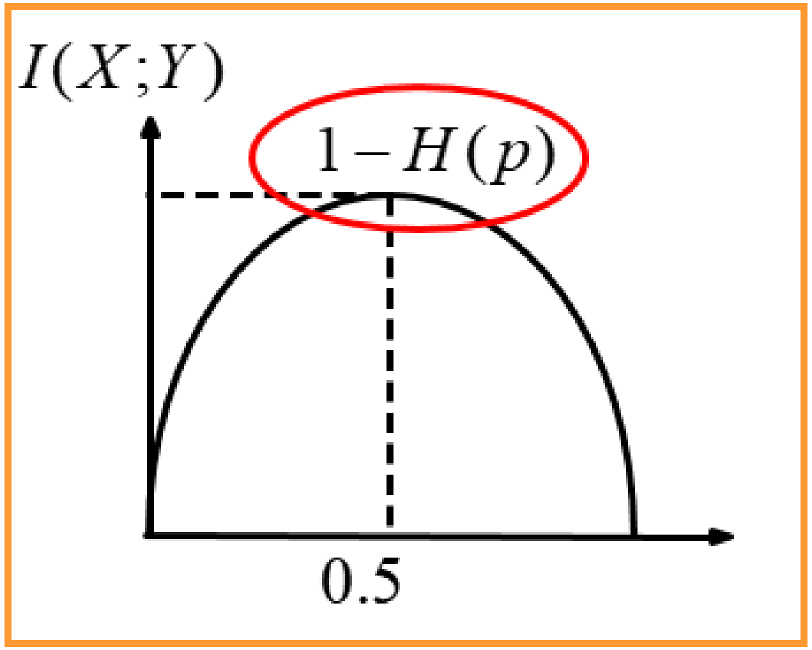
$$\boxed{C_t = \frac{1}{t}\max_{P(X)}\{I(X;Y)\} \quad （比特/秒）}$$

6.1、信道容量 - 例题 – 信道容量的计算

- 二元对称信道的信道容量

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \omega & \overline{\omega} \end{bmatrix} \qquad P_{Y|X} = \begin{bmatrix} \overline{p} & p \\ p & \overline{p} \end{bmatrix}$$

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

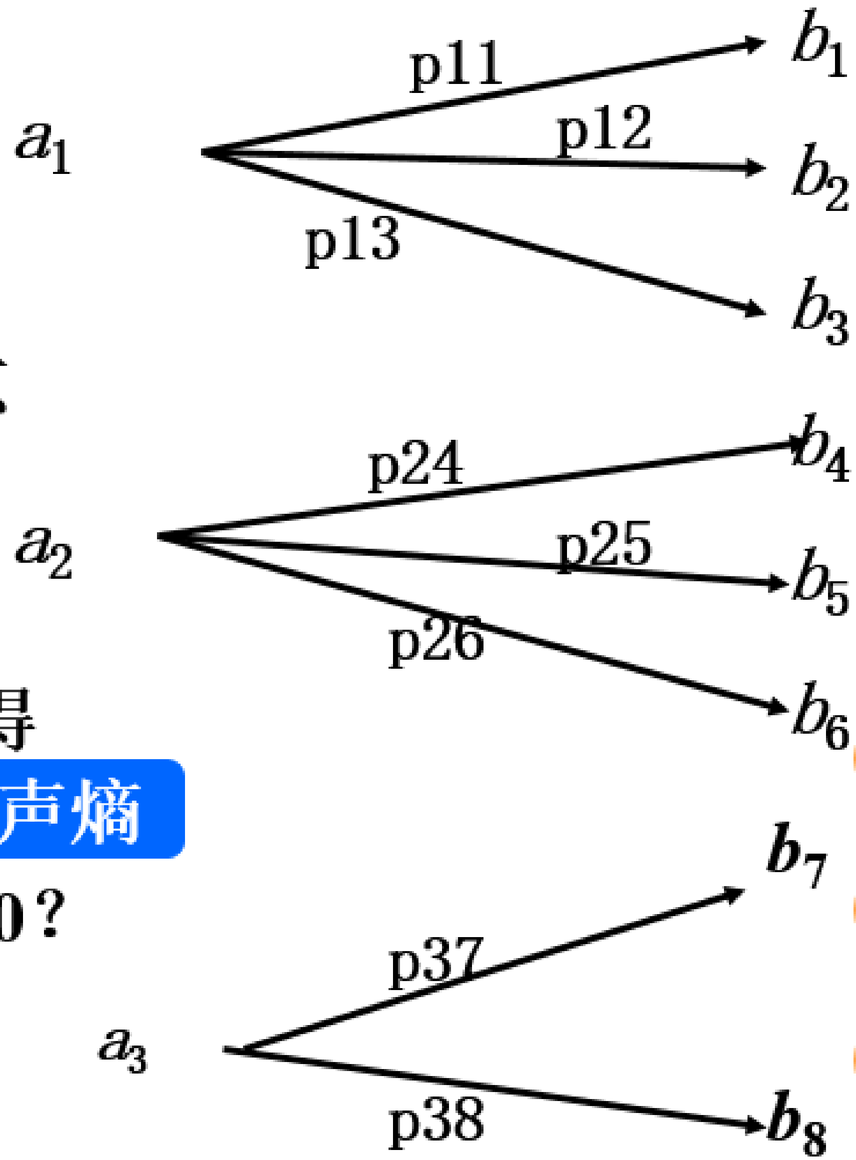


6.1、信道容量 – 核心问题

- 基本概念
  - 信道容量、最佳输入分布
  - 噪声熵、损失熵（疑义度）
- 主要内容
  - 特殊信道 - 输入/输出存在一定的确定关系
  - 离散对称信道、离散准对称信道
  - 无记忆扩展信道、独立并联信道
  - 信源和信道匹配

6.2、离散有噪无损信道 - 1

- 输出Y与输入X之间有着某种确定关系



1. 无损信道：有噪无损。一个输入对应多个互不相交的输出。
   1. 输入完全由输出决定；
   2. $p(y_j)\neq 0$，必存在x_i，使得 $p(x_i|y_j)=1$；
   3. $H(X|Y)$和$H(Y|X)$哪个为0？

噪声熵

损失熵

6.2、有噪无损信道特点

- 各输入符号对应的输出符号子集不相交

$$\mathbf{P}=\begin{bmatrix} p_{11} & p_{12} & p_{13} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{24} & p_{25} & p_{26} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_{37} & p_{38} \end{bmatrix}$$

$$H(X|Y)=0 \Rightarrow I(X;Y)=H(X)-H(X|Y)=H(X)$$

损失熵

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

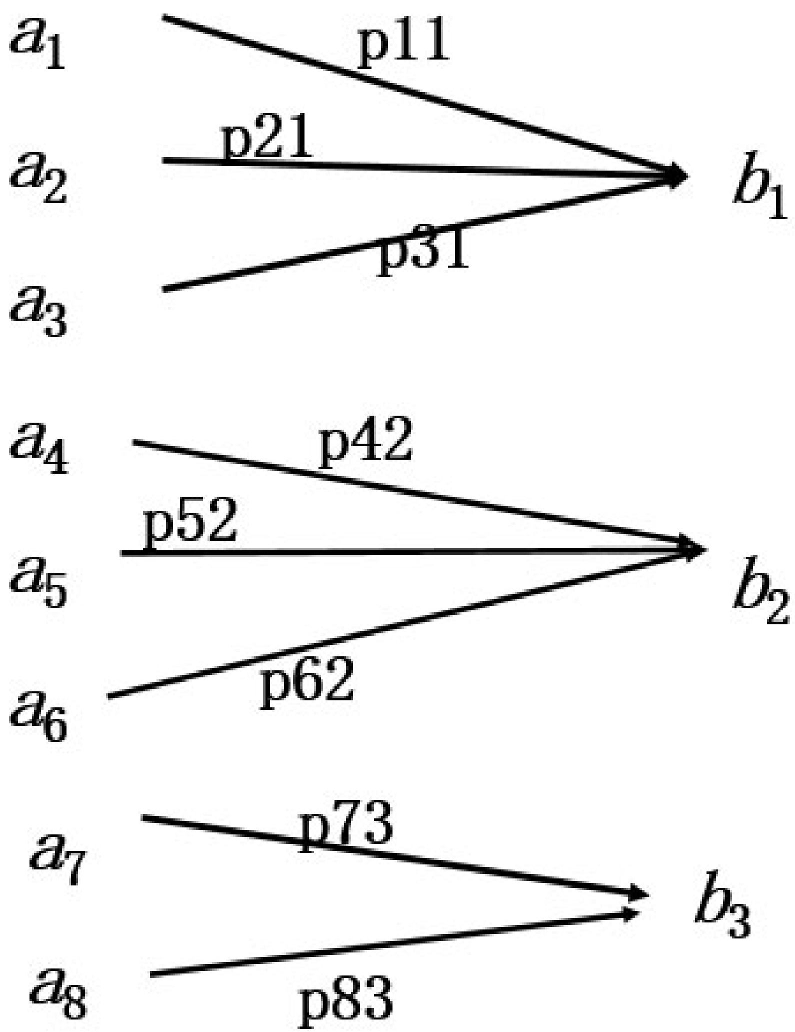
有噪无损信道容量 C=logr 取决于输入符号的个数！

6.2、离散无噪有损信道 - 2

2. 无噪信道：无噪有损
   - 多个输入对应一个输出。
   - 对所有x_i，必存在一个y_j，使得$p(y_j|x_i)=1$
   - $H(X|Y)$和$H(Y|X)$哪个为0？

损失熵

噪声熵



6.2、无噪有损信道特点

$$\mathbf{P}=\begin{bmatrix} p_{11} & 0 & 0 \\ p_{21} & 0 & 0 \\ p_{31} & 0 & 0 \\ 0 & p_{42} & 0 \\ 0 & p_{52} & 0 \\ 0 & p_{62} & 0 \\ 0 & 0 & p_{73} \\ 0 & 0 & p_{83} \end{bmatrix}$$

$$H(Y \mid X)=0$$

$$\Downarrow$$

$$I(X;Y)=H(Y)-H(Y \mid X)=H(Y)$$

$$\Downarrow$$

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

$$=\max_{p(x)} H(Y)=\log s$$

无噪有损信道容量
C＝*logs*

取决于输出符号的个数

6.2、离散无噪无损信道 - 3

3. 无噪无损信道

- 输入和输出一一对应
- 对所有x_i，必存在一个y_j，使得$p(y_j|x_i) = 1$
- $H(Y|X) = 0$;
- $H(X|Y) = 0$;

$$
\begin{array}{ccc}
a_1 & \xrightarrow{\;1\;} & b_1 \\
a_2 & \xrightarrow{\;1\;} & b_2 \\
a_3 & \xrightarrow{\;1\;} & b_3 \\
\vdots & & \vdots \\
a_r & \xrightarrow{\;1\;} & b_r
\end{array}
$$

$$p(b_j \mid a_i) = p(a_i \mid b_j) = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$$

6.2、无噪无损信道特点

- 信道矩阵为单位阵
- $r = s$

$$\mathbf{p} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & & \cdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$H(Y \mid X) = 0, H(X \mid Y) = 0$$

$$\Downarrow$$

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

$$\Downarrow$$

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

无噪无损信道容量 $C = \log r = \log s$

上堂课程主要内容复习

- 无记忆扩展信道 – 例题、作业
  - 信道矩阵、互信息、条件互信息
- 信道的组合
  - 级联信道、数据处理定理、等价信道
- 信道容量C $\boxed{C=\max\limits_{P(X)}\{I(X;Y)\} \quad （比特/符号）}$
  - 基本概念：平均互信息、信息传输率、信道容量、最佳输入分布、噪声熵、损失熵（疑义度）
  - C的求解：特殊信道的C求解

上堂课程主要内容复习

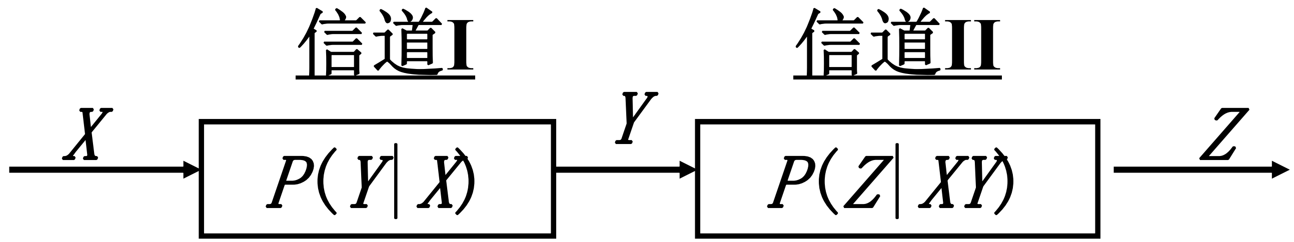
- 级联信道 $I(XY;Z)=\sum_{XYZ} p(xyz)\log\frac{p(z|xy)}{p(z)}$

$$I(XY;Z)\ge I(Y;Z) \qquad I(XY;Z)\ge I(X;Z)$$

- **X**、**Y**、**Z** 满足马尔可夫链时

$$I(X;Z)\le I(X;Y) \qquad I(X;Z)\le I(Y;Z)$$

- 数据处理定理 – 信息不增性原理



$$A=\{a_1,a_2,\cdots,a_r\} \qquad B=\{b_1,b_2,\cdots,b_s\} \qquad C=\{c_1,c_2,\cdots,c_l\}$$

$$p(b_j\mid a_i) \qquad p(c_k\mid a_ib_j)$$

6.3、离散对称信道 – 基本概念

- 离散输入对称信道（行对称信道）
  - 信道矩阵：每一行都是第一行元素的不同组合

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

- 离散输出对称信道（列对称信道）
  - 信道矩阵：每一列都是第一列元素的不同组合

$$\mathbf{P}=\begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$$

6.3、离散对称信道 – 概念（续）

- 对称信道
  - 每行都是第一行的排列
  - 并且，每列都是第一列的排列

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

行对称信道

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

对称信道

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

对称信道

6.3、离散对称信道 － 概念（续）

- 准对称信道
  - 不是对称信道。
  - 信道矩阵可以按列分为一些对称的子阵，即，对输出集Y的划分。
  - 划分子集只有一个 --〉对称信道。

$$
\mathbf{P} = \begin{bmatrix} 0.8 & \boxed{\begin{matrix} 0.1 \\ 0.1 \end{matrix}} & 0.1 \\ 0.1 & & 0.8 \end{bmatrix} \qquad \mathbf{P}_1 = \begin{bmatrix} 0.8 & 0.1 \\ 0.1 & 0.8 \end{bmatrix} \qquad \mathbf{P}_2 = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}
$$

6.3、离散对称信道 – 概念（续）

- 强对称矩阵（均匀矩阵）
  - $r = s$
  - 对于每个输入符号，正确传输概率都相等
  - 错误传输概率 p 均匀地分配到 r-1 个符号

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

$r \times r$ 阶对称矩阵

6.3、离散对称信道 – C 求解之引理

- 对于对称信道，当信道输入概率分布为等概分布时，输出概率分布必为等概分布。

证明 $p(a_i)=\frac{1}{r} \qquad i=1,\cdots,r$

$$p(b_j)=\sum_i p(a_i)p(b_j|a_i)=\frac{1}{r}\sum_i p(b_j|a_i)=$$

H_j 为信道矩阵第 j 列元素之和。而对称信道每一列是第一列的不同排列。因此 $H_1=H_2=\cdots=H_s$

$$p(b_1)=p(b_2)=\cdots=p(b_s)=\frac{1}{s}$$

对离散输出对称信道，此引理亦成立

6.3、离散对称信道 － 信道容量求解定理

- 对称信道 当信道输入概率分布为等概的情况下达到信道容量：

★★

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

其中 $p_1',p_2',\cdots,p_s'$ 是信道矩阵中的任意一行中的元素。

分析：
$$\log s=\max_{p(y)} H(Y)$$

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

6.3、离散对称信道 － 定理证明

$$H(Y|X) = -\sum_{i,j} p(x_i y_j)\log p(y_j|x_i)$$

$$= -\sum_i p(x_i)\underbrace{\sum_j p(y_j|x_i)\log p(y_j|x_i)}_{H(Y|x_i)}$$

$$= \sum_i p(x_i) H(p_1', p_2', \cdots, p_s') = H(p_1', p_2', \cdots, p_s')$$

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(p_1', p_2', \cdots, p_s')$$

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

$$= \max_{p(x)}\{H(Y) - H(p_1', p_2', \cdots, p_s')\}$$

$$= \max_{p(x)} H(Y) - H(p_1', p_2', \cdots, p_s')$$

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

1. 当输出*Y*达到等概分布时，***H*(*Y*)**达到最大值**log*s***。
2. 对称信道，输入等概分布才会使输出等概分布。

6.3、离散对称信道 －例题3.12

- 设某离散对称信道的信道矩阵如下，求解该信道的信道容量和对应的最佳输入分布。

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

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

最佳输入分布：信道输入符号为等概率分布

6.3、离散对称信道 － 推论

- 均匀信道的信道容量

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

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

$$=\log r-H(\bar{p},\frac{p}{r-1},\cdots,\frac{p}{r-1})$$

$$=\log r+\bar{p}\log\bar{p}+\frac{p}{r-1}\log\frac{p}{r-1}+\cdots+\frac{p}{r-1}\log\frac{p}{r-1}$$

$$=\log r+\bar{p}\log\bar{p}+p\log\frac{p}{r-1}$$

$$=\log r+\bar{p}\log\bar{p}+p\log p-p\log(r-1)$$

$$=\log r-H(p)-p\log(r-1)$$

6.4、一般离散信道 － 信道容量定理

信道容量定理 － 定理3.3

定理 $I(X;Y)$ 达到信道容量的充要条件是输入分布 $p(a_i)$ 满足以下条件：

$$p(a_i) \neq 0 \text{ 时，} \quad I(a_i;Y) = C$$

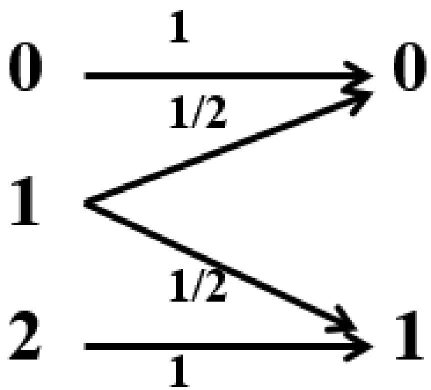
$$p(a_i) = 0 \text{ 时，} \quad I(a_i;Y) \leq C$$

其中 $I\left(a_i;Y\right) = \sum_j p(b_j \mid a_i) \log \frac{p(b_j \mid a_i)}{p(b_j)}$

某些特殊矩阵可以利用这个方法可以推导得到C。

6.4、信道容量定理应用 – 一般离散信道
例题3.13

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：

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

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

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

6.4、一般离散信道 – 例题3.13（续）

可知，上述结果满足定理的充要条件：

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

- 达到该信道容量的输入概率就是前面假设的输入分布，即： $\boldsymbol{p(0)=p(2)=1/2}$， $\boldsymbol{p(1)=0}$
- 信道容量为 $\boldsymbol{log2}$

6.4、离散准对称信道 －信道容量求解

- 准对称信道 信道输入等概率分布时达到信道容量：

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

  - $p_1', p_2', \cdots, p_s'$ 是信道矩阵中的任一行中的元素；$H(Y)$为信道输入等概时对应的输出熵。

- 说明：
  - 对于对称信道，此公式亦适用；
  - 也可通过“令输入等概，求解平均互信息”获得信道容量C。

6.4、离散准对称信道 – 信道容量求解分析

- 为什么是“信道输入等概率分布时”，准对称信道的传输能达到信道容量？
  - 设信道矩阵可划分为n个子矩阵，其中N_k是第k个子矩阵中行元素之和，M_k是第k个子矩阵中列元素之和。

$$I(x_i;Y)=\sum_Y p(y_j|x_i)\log\frac{p(y_j|x_i)}{p(y_j)}$$

$$=\underline{\sum_Y p(y_j|x_i)\log p(y_j|x_i)}-\sum_Y p(y_j|x_i)\log p(y_j)$$

输入等概 $p(y_j)=\frac{1}{r}*M_k$

行对称，与输入x取值无关

6.4、离散准对称信道 － 信道容量求解分析

根据定理，输入等概时，该信道达到信道容量！

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

分析：

信道矩阵中各行元素内容相同，则：

$$C = \max_{p(x)} \{H(Y)\} - H(p_1', p_2', \ldots, p_s')$$

$$p(y_j) = \frac{1}{r} * M_k \Rightarrow H(Y) = \sum_{Y} \frac{M_k}{r} \log \frac{r}{M_k}$$

$$H(Y) = \sum_{n} N_k \log \frac{r}{M_k} = \log r - \sum_{n} N_k \log M_k$$

$$C = \log r - H(p_1', p_2', \cdots, p_s') - \sum_{k=1}^{n} N_k \log M_k$$

6.4、离散准对称信道 – 信道容量求解

- 准对称信道 信道输入等概率分布时达到信道容量：

★
$$C = \log r - H(p_1', p_2', \cdots, p_s') - \sum_{k=1}^{n} N_k \log M_k$$

$$\mathbf{P} = \begin{bmatrix} 0.8 & \boxed{\begin{matrix} 0.1 \\ 0.1 \end{matrix}} & 0.1 \\ 0.1 & & 0.8 \end{bmatrix} \qquad \mathbf{P}_1 = \begin{bmatrix} 0.8 & 0.1 \\ 0.1 & 0.8 \end{bmatrix} \qquad \mathbf{P}_2 = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$$

- 设信道矩阵可划分为n个子矩阵，其中N_k是第k个子矩阵中行元素之和，M_k是第k个子矩阵中列元素之和。

6.4、离散无记忆N次扩展信道

一般离散无记忆信道 $I(\mathbf{X};\mathbf{Y}) \le \sum_{i=1}^{N} I(X_i;Y_i)$

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

离散无记忆N次扩展信道 $C_i = C$

$$C^N = \max_{p(x)} \{I(\mathbf{X};\mathbf{Y})\} = \max_{p(x)} \sum_{i=1}^{N} I(X_i;Y_i) = \sum_{i=1}^{N} C_i = NC$$

$$I(\mathbf{X};\mathbf{Y}) \le NC$$

输入信源无记忆，且序列中每一分量的分布各自达到最佳分布时，信道容量方可达到*NC*。

6.4、独立并联信道的信道容量

$X_1 \longrightarrow$ 信道1 $\longrightarrow Y_1$

$p(y_1 | x_1)$

$X_2 \longrightarrow$ 信道2 $\longrightarrow Y_2$

$p(y_2 | x_2)$

…

$X_N \longrightarrow$ 信道N $\longrightarrow Y_N$

$p(y_n | x_n)$

特点：每个信道的输出仅与本信道的输入有关。

当信道矩阵P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时，等号成立

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

6.5、信源和信道匹配

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。

$$\text{信道剩余度} = C - I(X;Y)$$

$$\text{相对剩余度} = [C - I(X;Y)] / C$$

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

$$\text{无损信道的相对剩余度} = 1 - H(X)/\log r$$

信源剩余度

6.5、信源和信道匹配 － 讨论

$$\boxed{1-I(X;Y)/C \quad 1-H_{\infty}/H_0}$$

- 提高无损信道传输率的研究 → 减少信源剩余度！
- 对于无损信道，可通过信源编码减少信源剩余度，使信息传输率趋于信道容量。
- 在信源的统计分析中，强调<u>“每个信源符号能提供多少信息量”</u>；在进入信道传输时，需要将信源符号变换成适合信道传输的信道符号（码符号），则问题转换为<u>“每个信道符号能传输多少信息量”</u>。
- 平均互信息 → 描述每个信道符号可以实际传输的信息量！

第三章 总结

- 信道的数学模型及其分类
- 互信息及其性质
- 离散无记忆信道
  - 熵、信道疑义度、平均互信息之间的关系
- 序列通过离散信道传输时
  - 无记忆扩展信道
- 信道的组合
  - 级联信道的平均互信息
  - 数据处理定理
- 信道容量 – 计算题类型

作业

- 1、2、4
- 7、8
- 9、10、15、16