信息论与编码原理

课程基本情况

- 专业基础课，48学时/3学分
- 先修课程
  - 概率论与数理统计，线性代数，高等数学
- 课程时间
  - 前4后2

教材

- 《信息论与编码》,傅祖芸,电子工业

- 参考教材
  - 《信息论 - 基础理论与应用》，傅祖芸，电子工业
  - 《信息论基础（原书第二版）》，Cover，机械
  - 《通信新读 – 从原理到应用》，陈小锋，机械

教学要求（安排）

- 课堂听讲
  - 主要教学环节
  - 课堂考勤；课堂秩序；
- 作业
  - 及时理解和掌握所学知识的主要手段
  - 作业纸；**抄写题目！**
- 答疑、其它交流方式
  - 微信

考核方法

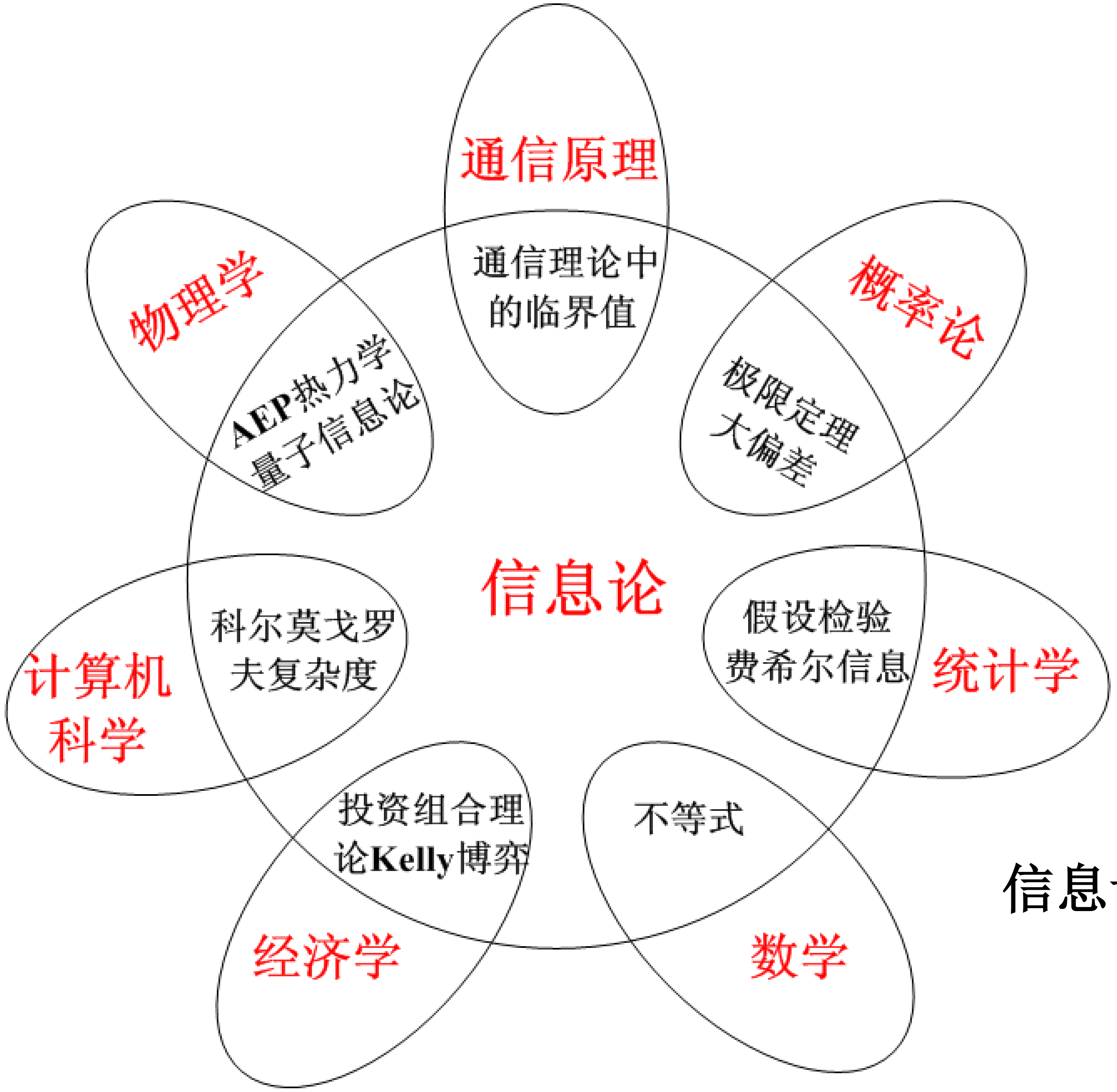
- 平时成绩 + 期末考试
  - 平时成绩：

    考勤&作业30%+小测
  - 期末考试：70%

第一章　　绪论

第一章　　绪论

- 信息的概念
- 信息论研究的对象、目的和内容★
- 信息论发展简史和信息科学
- 概率论知识的回顾 ★

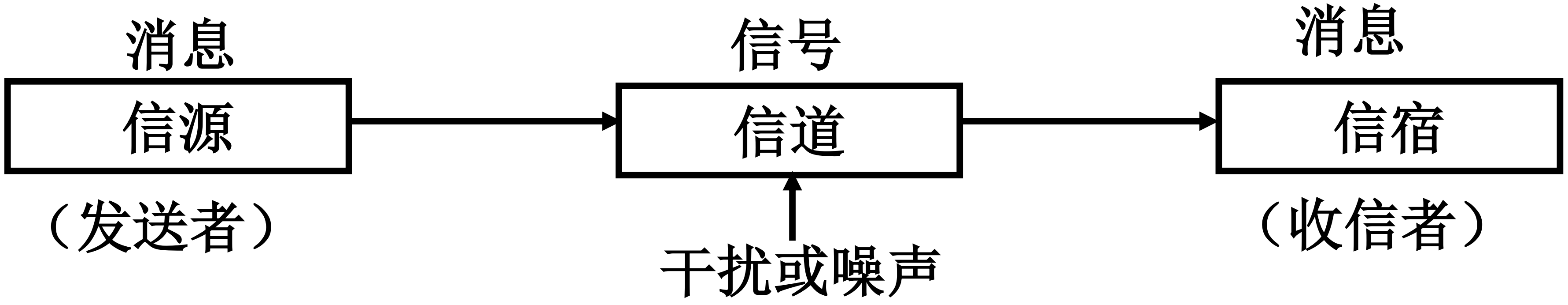

通信原理
通信理论中
的临界值
物理学
AEP热力学
量子信息论
概率论
极限定理
大偏差
信息论
计算机
科学
科尔莫戈罗
夫复杂度
假设检验
费希尔信息
统计学
投资组合理
论Kelly博弈
不等式
经济学
数学

信息论与其他学科关系

1.1 信息的概念

- 客观世界的三大要素
  - 物质
    - **Without materials nothing exists**
  - 能量
    - **Without energy nothing happens**
  - 信息
    - **Without information nothing makes sense**
  -

1.1 什么是信息？



点对点通信系统模型

通信的基本问题：

在一点精确地或近似地恢复另一点所选择的消息。

一点： place/time　e.g. …

传输和存储都属于通信问题！

1.1 什么是信息？

消息 → 信号 → 消息

信源（发送者） → 信道 → 信宿（收信者）

干扰或噪声 → 信道

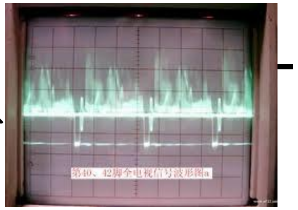
信息

⇧

消息

⇧

信号



- 消息：能被人的感觉器官所感知perception
  - 消息中包含信息，是信息的载体。是具体的，非物理的。声音、文字、图像、符号
- 信号：适合信道传输的物理量。
  - 可测量，可显示，可描述，携带消息，是消息的运载工具。电信号、光信号

1.1 什么是信息？

- 信息到底是什么？
  - 抽象☹
  - 能否定量定性分析？

1.1 什么是信息？

朗文双解：

Information: Facts or details that tell you something about a situation, person, event, etc.

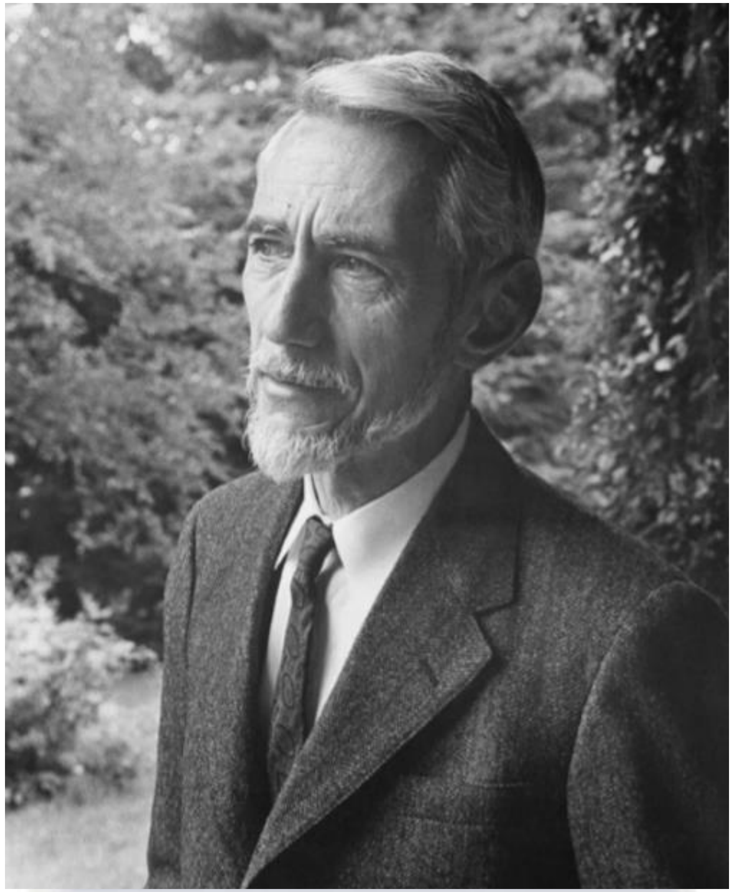
- 哈特莱－首先提出“信息”一词
  - 信息不同于消息
- 维纳（**N.Wiener**），*控制论*，**1948**
  - 信息：是区别于物质和能量的最基本概念
  - 与外界进行交换，混淆物质和能量
- 香农－信息论创始人
  - 通信过程消除了消息的不确定性，减少了未知

1.1 信息的概念- 香农信息

- 香农(C.E.Shannon) 1916~2001

  1948，通信的数学理论

  - 对信息作出了科学的定义，并进行了定量和定性的描述
  - 信息是事物运动状态或存在方式下不确定性的描述
  - 通信过程是一种消除不确定性的过程



Claude E. Shannon 1916-2001

1.1 什么是信息？- 对信息的理解

- 信息的本质
  - 信息被接收前，具有未知性，不确定性
- 数学语言
  - 不确定性就是：随机性、随机事件
- 数学工具
  - 概率论和随机过程来测度不确定性的大小
- 直观上讲
  - 不确定性大小←——→事先猜测某随机事件是否发生的难易程度

1.1 什么是信息？ - 例

- 一：明天太阳将从东方升起
- 二：一周内会下雨
- 三：今年海南会下雪
- 四：奖券问题

1.1 什么是信息？ - 信息的定义

- 不确定性的大小与事件发生的概率有关

- 不确定性是概率的函数

- 因此，信息量可以表示为概率的函数

1.1 什么是信息？ - 信息的定义

- 考虑可能输出为x_i，$i=1,2,\ldots n$的离散随机变量X。则事件$X=x_i$的自信息（**Self-Information**）表示为：

$$\boxed{I(x_i) = -\log p(x_i) = \log \frac{1}{p(x_i)}}$$

1.1 什么是信息？ - 香农信息的特点

- 优点
  - 有明确的数学表达式，定量化
  - 与人们直观理解的信息含义比较一致
  - 排除信息一词主观上的含义，同一消息对任何收信者，所得信息量相同
- 局限性
  - 假设事物状态可以用一个以经典集合论为基础的概率模型来描述
  - 没有考虑收信者的主观特性和主观意义

第一章 绪论

- 信息的概念
- 信息论研究的对象★、目的和内容
- 信息论发展简史和信息科学
- 概率论知识的回顾 ★

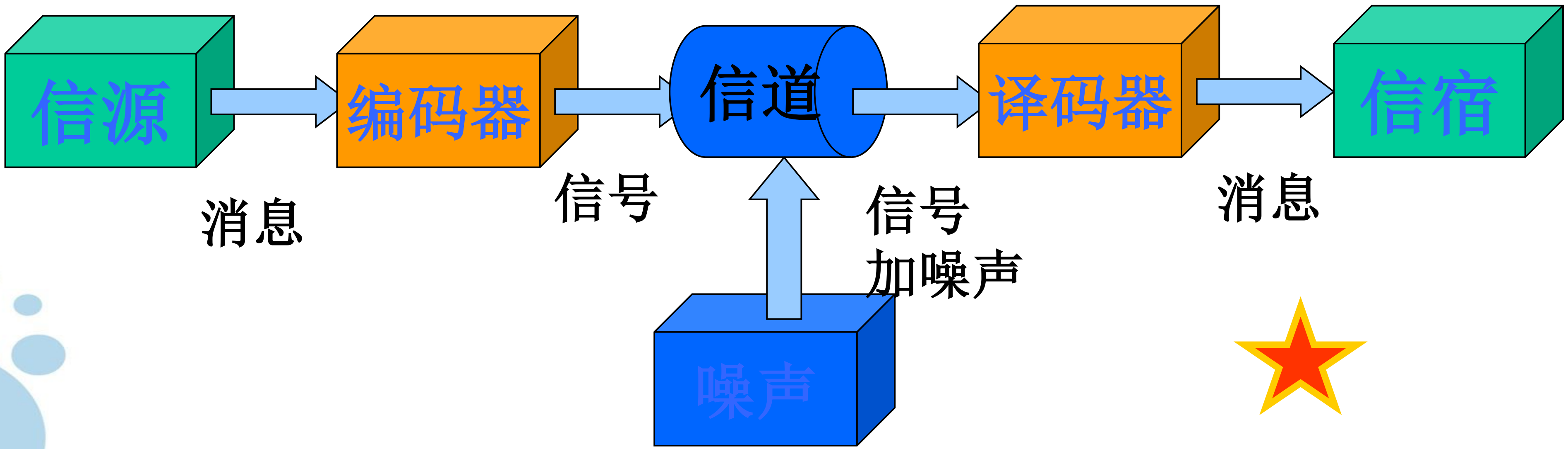
1.2 通信系统模型—信息论研究对象

- 信息：事物 运动状态 或 存在方式下不确定性的描述。 “传递”与“获取”

“通信的基本问题就是在一点重新准确地或近似地再现另一点所选择的消息”。

-----*香农*

- 通信系统形式上传输的是消息，实质上传输的是信息
- 研究信息传输和处理的共同规律



1.2 通信系统模型 – 编码器

- 功能：将消息变成适合信道传输的物理量
- 编码器包括：
  - 信源编码器 – 提高通信系统的有效性(去冗余)
  - 信道编码器 – 提高信息传输的可靠性(加冗余)
  - 调制器 – 变成适合信道传输要求的信号

举例

1.2 通信系统模型 – 编码器

– 有效性efficiency (去冗余) --信源编码器（调制器）

信息传输应该尽量快，高传输码率

手段：压缩、调制

–可靠性reliability（加冗余）--信道编码器

传输差错尽量少，低误码率

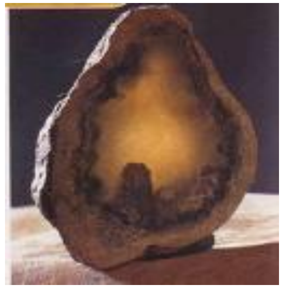
– 安全性

未经授权，无法获取。

1.2 通信系统模型 – 编码器

- 信源编码去除冗余度，信道编码却加上冗余度，为什么要这么做？

玉石原石—>玉石 —> 包装 —> 运输



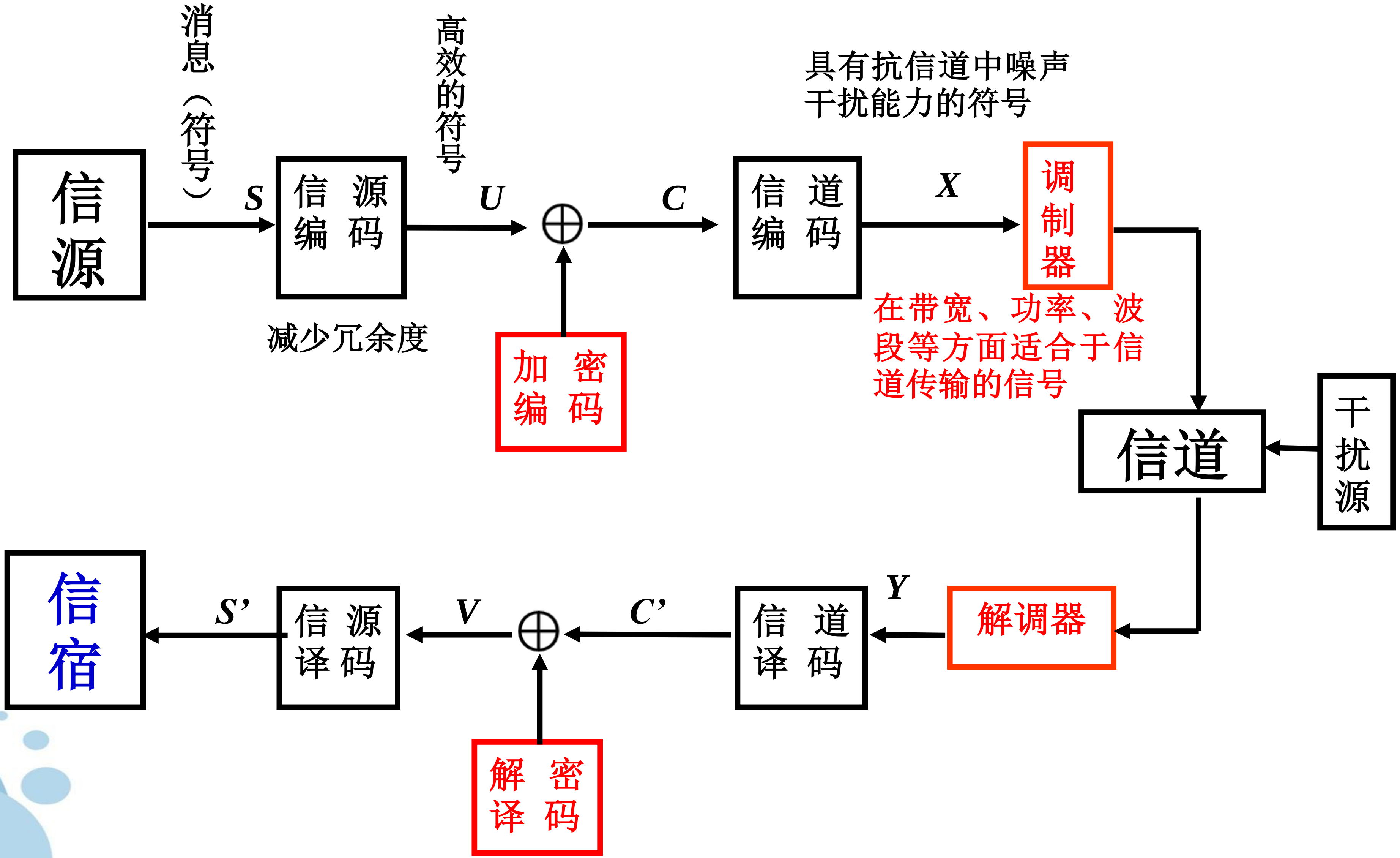
信源

信源编码

信道编码

调制 信道

1.2 通信系统模型 – 总结


消息（符号）
高效的符号
具有抗信道中噪声干扰能力的符号
信源
S
信 源 编 码
U
C
信 道 编 码
X
调制器
减少冗余度
加 密 编 码
在带宽、功率、波段等方面适合于信道传输的信号
信道
干扰源
信宿
S'
信 源 译 码
V
C'
信 道 译 码
Y
解调器
解 密 译 码

1.2 信息论的研究内容 – 特点

- 以概率论和随机过程为基本研究工具
- 研究的是通信系统的整个过程，而不是单个环节，并以编、译码器为重点
- 关心的是*最优系统的性能*和怎样达到这个性能（并不具体涉及系统—通信原理）
- 不研究信宿

信息论关心的是系统的理论极限和潜能。

“我们能对数据进行压缩的极限是什么？”

“我们能够达到的最佳纠错性能是什么？”

编码理论所关心的是创建实用的编码和译码系统。

信息论中的许多结果都相当微妙，以至于该学科的专家也不时地发现自己对某个结果认识不足，甚至是不正确的。

-----杨伟豪（香港中文大学）

第一章 绪论

- 信息的概念
- 信息论研究的对象、目的和内容
- 信息论发展简史和信息科学（课下阅读）
- 概率论知识的回顾

练习

1、已知随机变量的联合概率分布$p(x_i y_j)$如下所示，
用概率空间：表示P_{XY}，求P_X，P_Y，$P_{X|Y}$，$P_{Y|X}$

<table>
<tr><td colspan="2" rowspan="2">$p(x_i y_j)$</td><td colspan="3">y_j</td></tr>
<tr><td>1</td><td>2</td><td>3</td></tr>
<tr><td rowspan="3">x_i</td><td>1</td><td>1/4</td><td>1/18</td><td>0</td></tr>
<tr><td>2</td><td>1/18</td><td>1/3</td><td>1/18</td></tr>
<tr><td>3</td><td>0</td><td>1/18</td><td>7/36</td></tr>
</table>

2. 某年级有甲，乙，丙3个班级，各班人数分别占年级总人数的1/4，1/3，5/12。已知甲，乙，丙3个班级中集邮人数分别占该班总人数的1/2，1/4，1/5

- （1）事件“某人为集邮者”的概率；
- （2）事件“某人既为集邮者，又属于乙班” 的概率。
- （3）事件 “某集邮者是乙班的” 的概率.

概率论知识的回顾

预备知识

- 符号约定
  - 大写字母 X, Y, Z 等表示随机变量
  - 小写字母 $x_i, y_j, x_1, \ldots, x_n, y_1, \ldots, y_m$ 表示随机变量的具体取值，即随机事件

预备知识 – 概率空间

- **离散随机变量*X*的概率空间为：**

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ p(x_1) & p(x_2) & \cdots & p(x_n) \end{bmatrix} \quad 0 \le p(x_i) \le 1, \quad \sum_{i=1}^{n} p(x_i) = 1$$

$$\begin{aligned} P_X &= \begin{bmatrix} p(x_1) & p(x_2) & \cdots & p(x_n) \end{bmatrix} \\ &= \begin{bmatrix} p_{x_1} & p_{x_2} & \cdots & p_{x_m} \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix} \end{aligned}$$

预备知识 – 概率空间

- **离散随机变量*Y*的概率空间为：**

$$\begin{bmatrix} Y \\ P(Y) \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & \cdots & y_m \\ p(y_1) & p(y_2) & \cdots & p(y_m) \end{bmatrix} \quad 0 \le p(y_j) \le 1, \quad \sum_{j=1}^{m} p(y_j) = 1$$

$$P_Y = \begin{bmatrix} p(y_1) & p(y_2) & \cdots & p(y_m) \end{bmatrix}$$

$$= \begin{bmatrix} p_{y_1} & p_{y_2} & \cdots & p_{y_m} \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & \cdots & p_m \end{bmatrix}$$

预备知识 – 联合概率和条件概率

- 离散随机变量的联合概率空间为

二维矩阵表示

$$\begin{bmatrix} XY \\ P(XY) \end{bmatrix} = \begin{bmatrix} x_1y_1 & \cdots & x_iy_j & \cdots & x_ny_m \\ p(x_1y_1) & \cdots & p(x_iy_j) & \cdots & p(x_ny_m) \end{bmatrix} \quad 0 \le p(x_iy_j) \le 1, \qquad \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_iy_j) = 1$$

$$P_{XY} = \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \begin{bmatrix} p(x_1y_1) & p(x_1y_2) & \cdots & p(x_1y_m) \\ p(x_2y_1) & p(x_2y_2) & \cdots & p(x_2y_m) \\ \vdots & \vdots & \vdots & \vdots \\ p(x_ny_1) & p(x_ny_2) & \cdots & p(x_ny_m) \end{bmatrix}$$

预备知识 – 联合概率和条件概率

- 全概率公式（**求边缘概率**）

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

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

预备知识 – 联合概率和条件概率

- **贝叶斯(Bayes)公式:**

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

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

- 其它常用概率关系:

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

$$P_{X|Y} = \begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_m \end{array} \begin{bmatrix} p(x_1 \mid y_1) & p(x_2 \mid y_1) & \cdots & p(x_n \mid y_1) \\ p(x_1 \mid y_2) & p(x_2 \mid y_2) & \cdots & p(x_n \mid y_2) \\ \vdots & \vdots & \vdots & \vdots \\ p(x_1 \mid y_m) & p(x_2 \mid y_m) & \cdots & p(x_n \mid y_m) \end{bmatrix}$$

$$P_{Y|X} = \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \begin{bmatrix} p(y_1 \mid x_1) & p(y_2 \mid x_1) & \cdots & p(y_m \mid x_1) \\ p(y_1 \mid x_2) & p(y_2 \mid x_2) & \cdots & p(y_m \mid x_2) \\ \vdots & \vdots & \vdots & \vdots \\ p(y_1 \mid x_n) & p(y_2 \mid x_n) & \cdots & p(y_m \mid x_n) \end{bmatrix}$$

$$
\begin{array}{ccc}
 & & \begin{array}{ccc} y1 & \quad y2 & \quad y3 \end{array} \\
P_{Y|X} & = & \begin{array}{c} x1 \\ x2 \end{array} \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/3 & 2/3 \end{bmatrix}
\end{array}
\qquad
P_{Y|X} = \begin{bmatrix} p(y_1|x_1) & p(y_2|x_1) & p(y_3|x_1) \\ p(y_1|x_2) & p(y_2|x_2) & p(y_3|x_2) \end{bmatrix}_{2\times 3}
$$

$$
P_X = \begin{bmatrix} 1/4 & 3/4 \end{bmatrix} \qquad P_X = \begin{bmatrix} p(x_1) & p(x_2) \end{bmatrix}_{1\times 2}
$$

$$
P_X P_{Y|X} = \begin{bmatrix} [p(x_1)p(y_1|x_1) + p(x_2)p(y_1|x_2)] & \cdots & \cdots \end{bmatrix}
$$

$$
= \begin{bmatrix} p(y_1) & p(y_2) & p(y_3) \end{bmatrix} = P_Y
$$

矩阵：$P_Y = P_X P_{Y|X}$

$$
= \begin{bmatrix} \frac{1}{4}\times\frac{1}{2}+\frac{3}{4}\times 0 & \frac{1}{4}\times\frac{1}{2}+\frac{3}{4}\times\frac{1}{3} & \frac{1}{4}\times 0+\frac{3}{4}\times\frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{bmatrix}
$$

$$
p(x_1)p(y_1|x_1) + p(x_2)p(y_1|x_2) = p(y_1)
$$

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

$$P_X = \begin{bmatrix} 1/4 & 3/4 \end{bmatrix} \qquad P_{Y|X} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

$$P_X = \begin{bmatrix} p(x_1) & p(x_2) \end{bmatrix}_{1\times 2} \qquad P_{Y|X} = \begin{bmatrix} p(y_1|x_1) & p(y_2|x_1) & p(y_3|x_1) \\ p(y_1|x_2) & p(y_2|x_2) & p(y_3|x_2) \end{bmatrix}_{2\times 3}$$

$$P_{XY} = \begin{bmatrix} p(x_1y_1) & p(x_1y_2) & p(x_1y_3) \\ p(x_2y_1) & p(x_2y_2) & p(x_2y_3) \end{bmatrix}$$

$$= \begin{bmatrix} p(x_1)p(y_1|x_1) & p(x_1)p(y_2|x_1) & p(x_1)p(y_3|x_1) \\ p(x_2)p(y_1|x_2) & p(x_2)p(y_2|x_2) & p(x_2)p(y_3|x_2) \end{bmatrix}_{2\times 3}$$

$$= \begin{bmatrix} \frac{1}{4}\times\frac{1}{2} & \frac{1}{4}\times\frac{1}{2} & \frac{1}{4}\times 0 \\ \frac{3}{4}\times 0 & \frac{3}{4}\times\frac{1}{3} & \frac{3}{4}\times\frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1/8 & 1/8 & 0 \\ 0 & 1/4 & 1/2 \end{bmatrix}$$

3. 某校入学考试中有25%考生被录取。被录取的考生中有50%来自本市，而落榜考生中有25%来自本市。所有本市的考生都学过英语，外地考生中录取和未录取的都只有50%学过英语。问：

- （1）“某学生学过英语”该事件发生的概率
- （2）“学过英语的某学生被录取”该事件发生的概率
- （3）考生地域状态已知时英语学习状态的概率分布。

X表示某生是否被录取
Y表示某生是否来自本市
Z表示某生是否学过英语

1） 求？
2） 求？
3） 求？

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} a_1 = \text{录取} & a_2 = \text{落榜} \\ 1/4 & 3/4 \end{bmatrix} \qquad \begin{bmatrix} Y \\ P(Y) \end{bmatrix} = \begin{bmatrix} b_1 = \text{本市} & b_2 = \text{外市} \\ & \end{bmatrix}$$

$$\begin{bmatrix} Z \\ P(Z) \end{bmatrix} = \begin{bmatrix} c_1 = \text{学过英语} & c_2 = \text{没学过} \\ & \end{bmatrix}$$

$$P_X = \begin{array}{cc} a1 & a2 \end{array} \quad \begin{bmatrix} 0.25 & 0.75 \end{bmatrix}$$

$$P_{Y|X} = \begin{array}{c|c} & b1 \quad b2 \\ \begin{array}{c} a1 \\ a2 \end{array} & \begin{bmatrix} 0.5 & \\ 0.25 & \end{bmatrix} \end{array}$$

$$P_{Z|Y} = \begin{array}{c|c} & c1 \quad c2 \\ \begin{array}{c} b1 \\ b2 \end{array} & \begin{bmatrix} 1 & \\ 0.5 & \end{bmatrix} \end{array}$$

$$P_{Z|XY} = \begin{array}{c|c} & c1 \quad c2 \\ \begin{array}{c} a1b1 \\ a1b2 \\ a2b1 \\ a2b2 \end{array} & \begin{bmatrix} 1 & \\ 0.5 & \\ 1 & \\ 0.5 & \end{bmatrix} \end{array}$$

（1）“某学生学过英语”这一事件发生的概率***p(c1)*** **21/32**
（2）“学过英语的某学生被录取”这一事件发生的概率***p(a1|c1)*** **2/7**
（3）考生地域状态已知时英语学习状态的概率分布。$\boldsymbol{P_{Z|Y}}$

$$P_{Z|Y} = \begin{array}{c|c} & c1 \quad c2 \\ \begin{array}{c} b1 \\ b2 \end{array} & \begin{bmatrix} 1 & \mathbf{0} \\ 0.5 & \mathbf{0.5} \end{bmatrix} \end{array}$$

某校入学考试中有25%考生被录取。被录取的考生中有50%来自本市，而落榜考生中有25%来自本市。所有本市的考生都学过英语，外地考生中录取和未录取的都只有50%学过英语。

总结

- 从问题中理清随机变量
- 分析随机变量的概率空间
  - 先清楚随机变量的取值，再分析概率
- 在分析多个随机变量或随机变量序列时
  - 找到变量之间的条件概率/联合概率分布
  - 是否统计独立？
  - 诀窍：找到联合概率分布（往往通过条件概率），一切迎刃而解！

作业

1.　某城市天气情况与气象预报分别视为包含{雨，无雨}的随机变量X和Y，且X与Y的联合概率为:

P(雨，雨)=1/8,
P(雨，无雨)=1/16,
P(无雨，雨)=3/16,
P(无雨，无雨)=10/16,
求:
1）求天气预报的准确率
2） $P_{X|Y}$ (*矩阵表示*)

2. 某年级有甲，乙，丙3个班级，各班人数分别占年级总人数的1/4，1/3，5/12。已知甲，乙，丙3个班级中集邮人数分别占该班总人数的1/2，1/4，1/5

- （1）事件“某人为集邮者”的概率；
- （2）事件“某人既为集邮者，又属于乙班” 的概率。
- （3）事件 “某集邮者是乙班的” 的概率.

3. 某校入学考试中有**25%**考生被录取。被录取的考生中有**50%**来自本市，而落榜考生中有**25%**来自本市。所有本市的考生都学过英语，外地考生中录取和未录取的都只有**50%**学过英语。问:

- （**1**）“某学生学过英语”这一事件发生的概率
- （**2**）“学过英语的某学生被录取”这一事件发生的概率
- （**3**）考生地域状态已知时英语学习状态的概率分布。