概率论与数理统计

王笑尘

北京邮电大学网络空间安全学院

wxiaochen@bupt.edu.cn

主要内容

- 课程介绍
- 课程考核及安排
- 第一章 概率论的基本概念

什么是概率论与数理统计？

- **必然现象中的确定性规律：**
  - **1+1=2**
  - 在标准大气压下，纯净水的沸点为**100**摄氏度

- **自然界和社会生活中还存在大量不确定现象：**
  - 人的寿命
  - 天气现象—明天是否下雨
  - 扔硬币的正反面

什么是概率论与数理统计?

■ 虽然存在不确定性，但还是有某些规律：

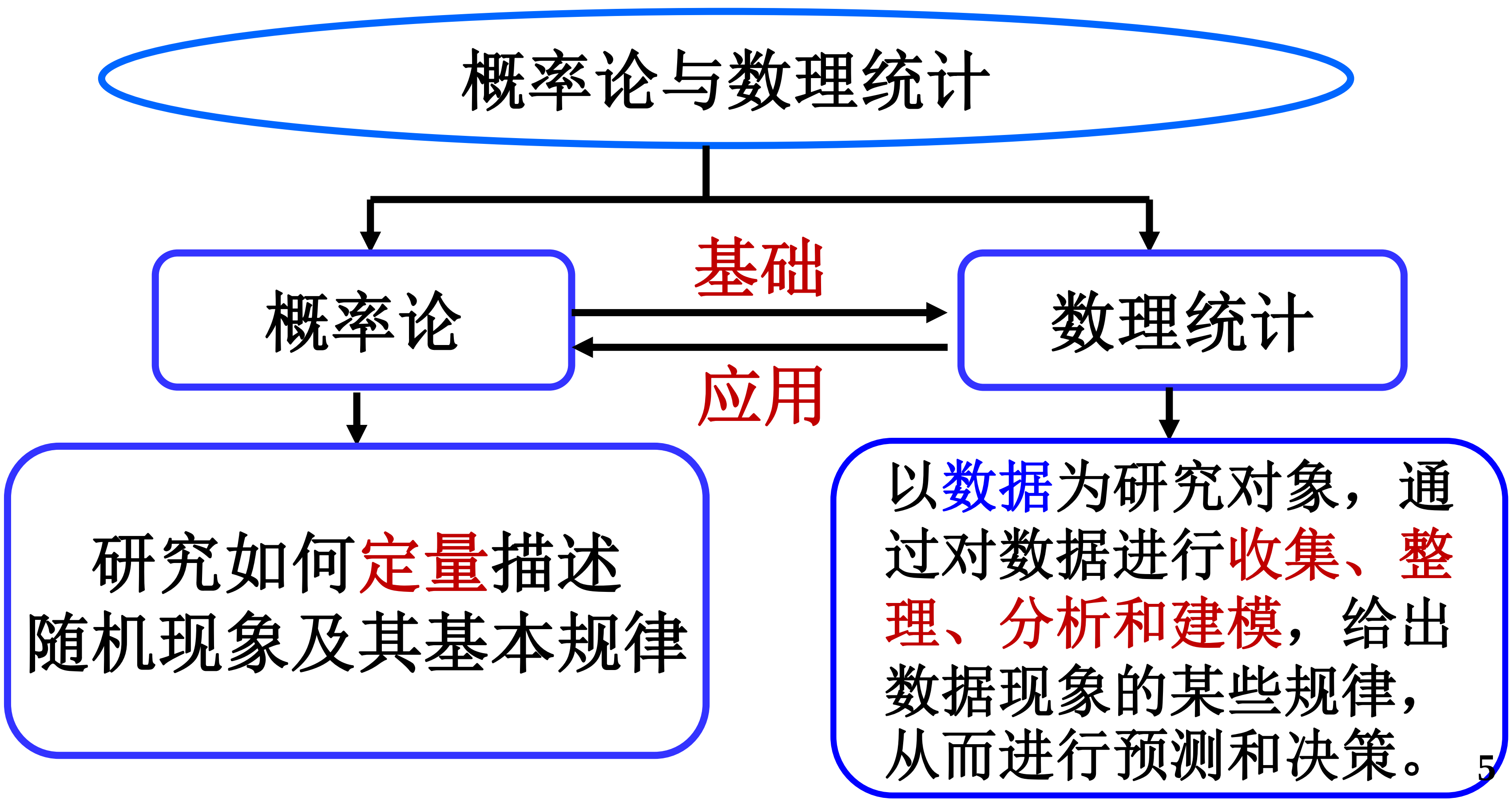
➢ 概率学家皮尔逊做扔硬币的试验

$$
\textbf{24000次}\begin{cases} 12012\text{次正面，占}0.5005 \approx \dfrac{1}{2} \\ 11988\text{次反面，占}0.4995 \approx \dfrac{1}{2} \end{cases}
$$

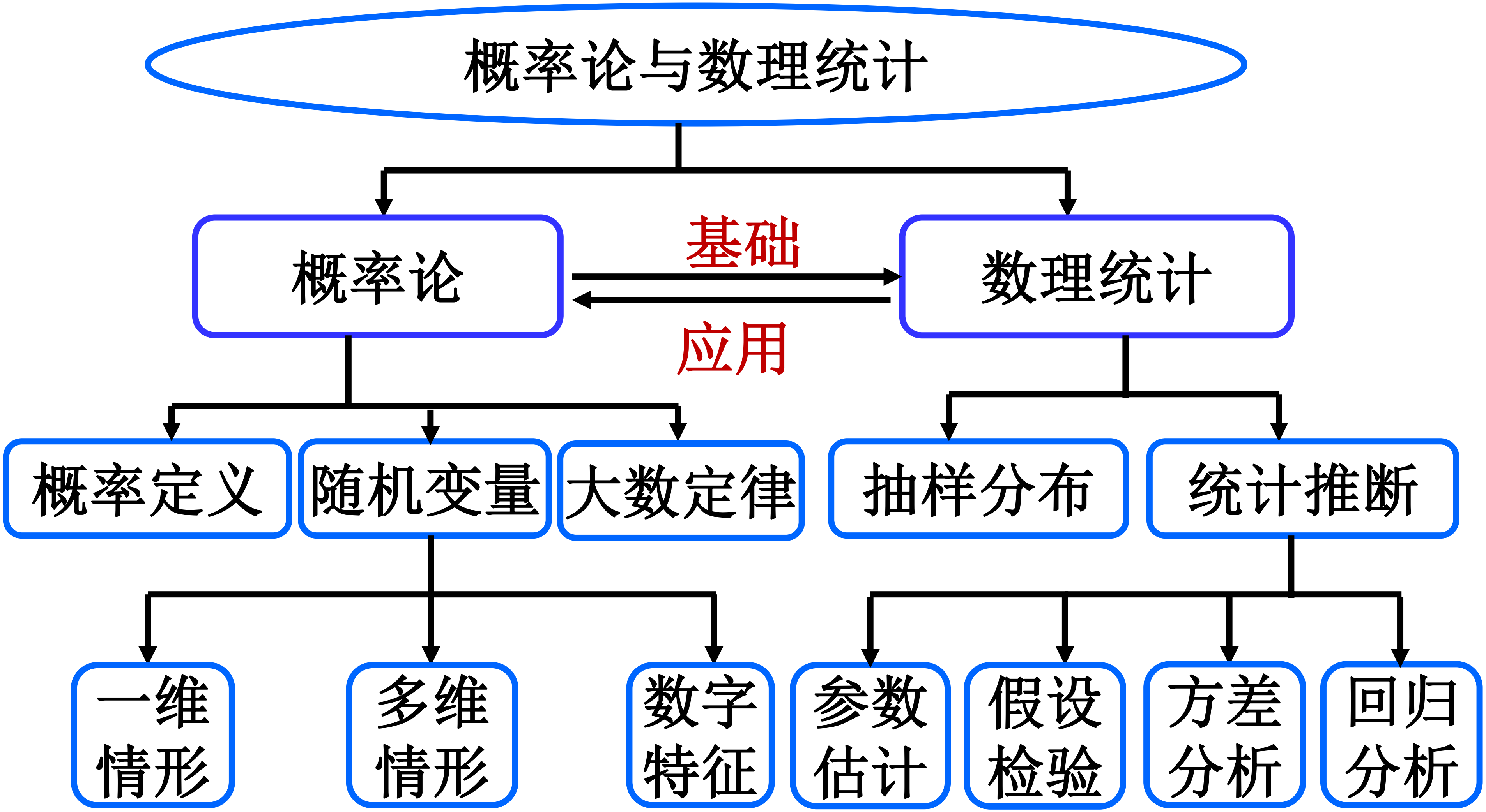
■ 在个别试验中其结果呈现不确定性，在大量重复试验中其结果又具有统计规律性的现象，称为随机现象。

什么是概率论与数理统计？

- 概率论与数理统计是研究和揭示随机现象统计规律性的一门数学学科。



知识结构图


概率论与数理统计
概率论
数理统计
基础
应用
概率定义
随机变量
大数定律
抽样分布
统计推断
一维情形
多维情形
数字特征
参数估计
假设检验
方差分析
回归分析

怎样学习《概率论与数理统计》课程

- 学思想。如何看待和处理随机规律性，是其它学科中没有的。例如，以比较各种事件出现的可能性的大小进行决策的思想。
- 学方法。定量描述随机现象及其规律的方法，收集、整理、分析数据，从而建立统计模型的方法。
- 学应用。尽可能多地了解各种概念的背景、各种方法和模型的实际应用。不仅要学课程中提及的，也要自己收集、寻找各种实例，从而解决实际问题。

主要内容

- 课程介绍
- 课程考核及安排
- 第一章 概率论的基本概念

课程考核

- **课堂要求：** 1、关闭手机或调静音；2、不要迟到
- **课程考核：**
  - **考勤+平时作业（占20分）：** 随机考勤；作业来自教材课后习题，通过教学云平台提交，写上姓名、班级、学号和页码（如1/5）；作业不得抄袭，有创新性思路或解决方案则另加分。

课程考核

- ■ 课程考核：
- ➢ 期中论文（占20分）：在以下两种题型之一完成论文1篇。题型一：应用案例。结合教学知识点，给出贴近生活或者本专业学习的应用案例及解答。题型二：题型归纳。通过几个相关联的题目，总结某类题型的解题思路。
- ➢ 期末考试（占60分）：统一执行闭卷书面考试。

课程安排

教学内容	学时	周次
第一章 概率论的基本概念	6学时	3-4周（第一章作业1次）
第二章 随机变量及其分布函数	6学时	5-6周（第二章作业1次）
第三章 多维随机变量及其分布	6学时	7-8周（第三章作业1次）
第四章 随机变量的数字特征	6学时	9-10周（第四章作业1次）
第五章 大数定律及中心极限定理	4学时	10-11周（基于前五章，准备期中论文）
第六章 数理统计部分（样本及抽样分布）	4学时	11-12周
第七章 参数估计	6学时	12-13周（第五、六、七章作业1次）
第八章 假设检验	8学时	14-15周
第九章 方差分析	6学时	16-17周（第八、九章作业1次）
第九章 回归分析	6学时	17-18周

课程安排

教学内容	知识点
第一章 概率论的基本概念	随机事件及其运算，事件的概率及其性质，条件概率，事件的独立性
第二章 随机变量及其分布函数	离散型随机变量及其分布律，连续型随机变量及其概率密度，随机变量函数及其分布
第三章 多维随机变量及其分布	多维随机变量及其联合分布，边缘分布，条件分布，随机变量的独立性，多维随机变量函数及其分布，n维随机变量简介
第四章 随机变量的数字特征	数学期望，方差，协方差与相关系数，矩

课程安排

教学内容	知识点
第五章 大数定律及中心极限定理	大数定律，中心极限定理
第六章 数理统计部分（样本及抽样分布）	总体、样本统计量，χ^2分布、t分布、F分布，抽样分布，顺序统计量
第七章 参数估计	矩估计及极大似然估计，估计量的评选标准，区间估计
第八章 假设检验	假设检验问题与基本概念，正态总体期望与方差假设检验，曲线拟合，齐一性及独立性的χ^2检验法，偏度、峰度检验法，秩和检验法
第九章 方差分析及回归分析	单因素方差分析，双因素方差分析，一元线性回归，多元线性回归

课程资源

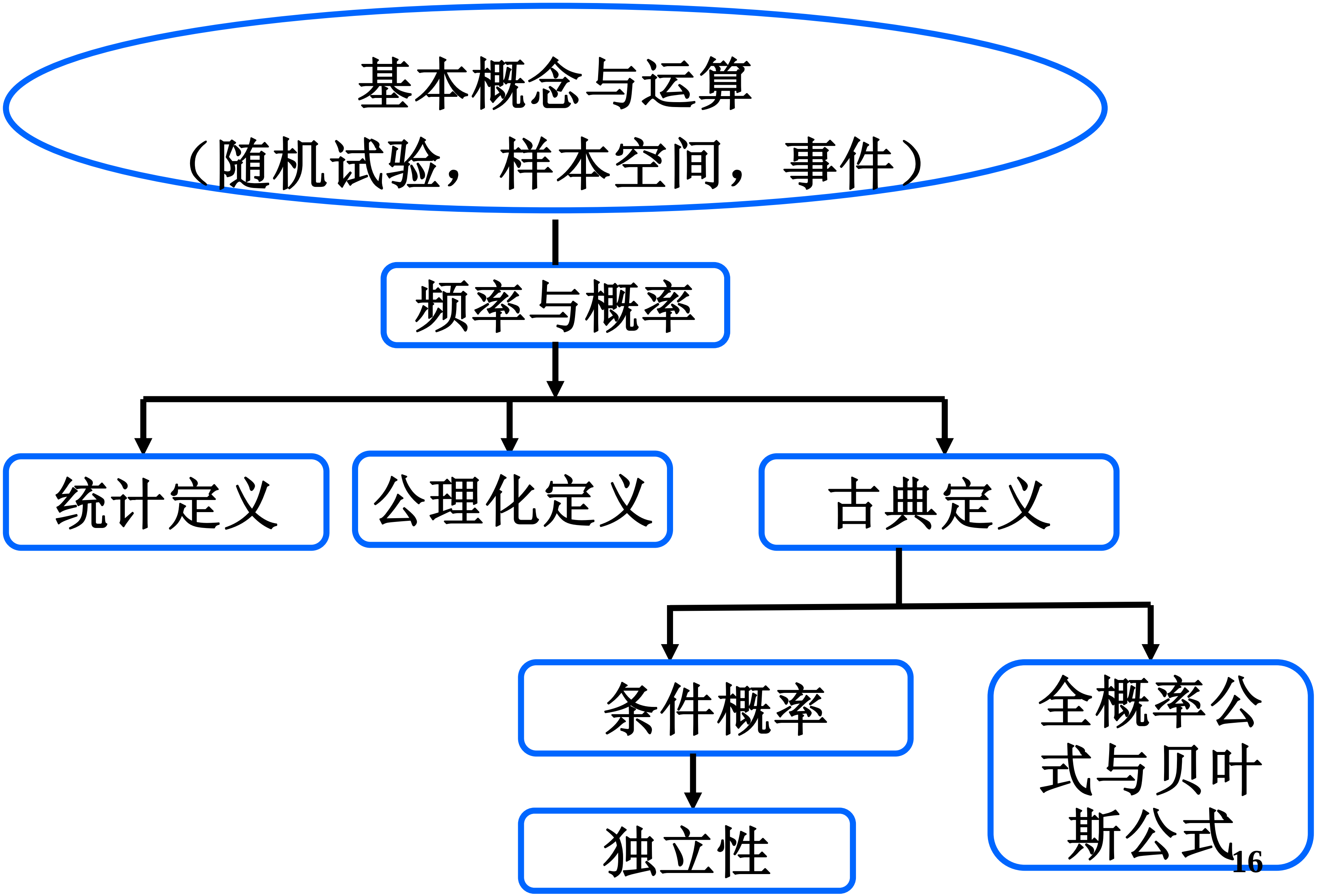
《概率论与数理统计》第五版，盛骤等，高等教育出版社，2019.

线上：中国大学MOOC平台

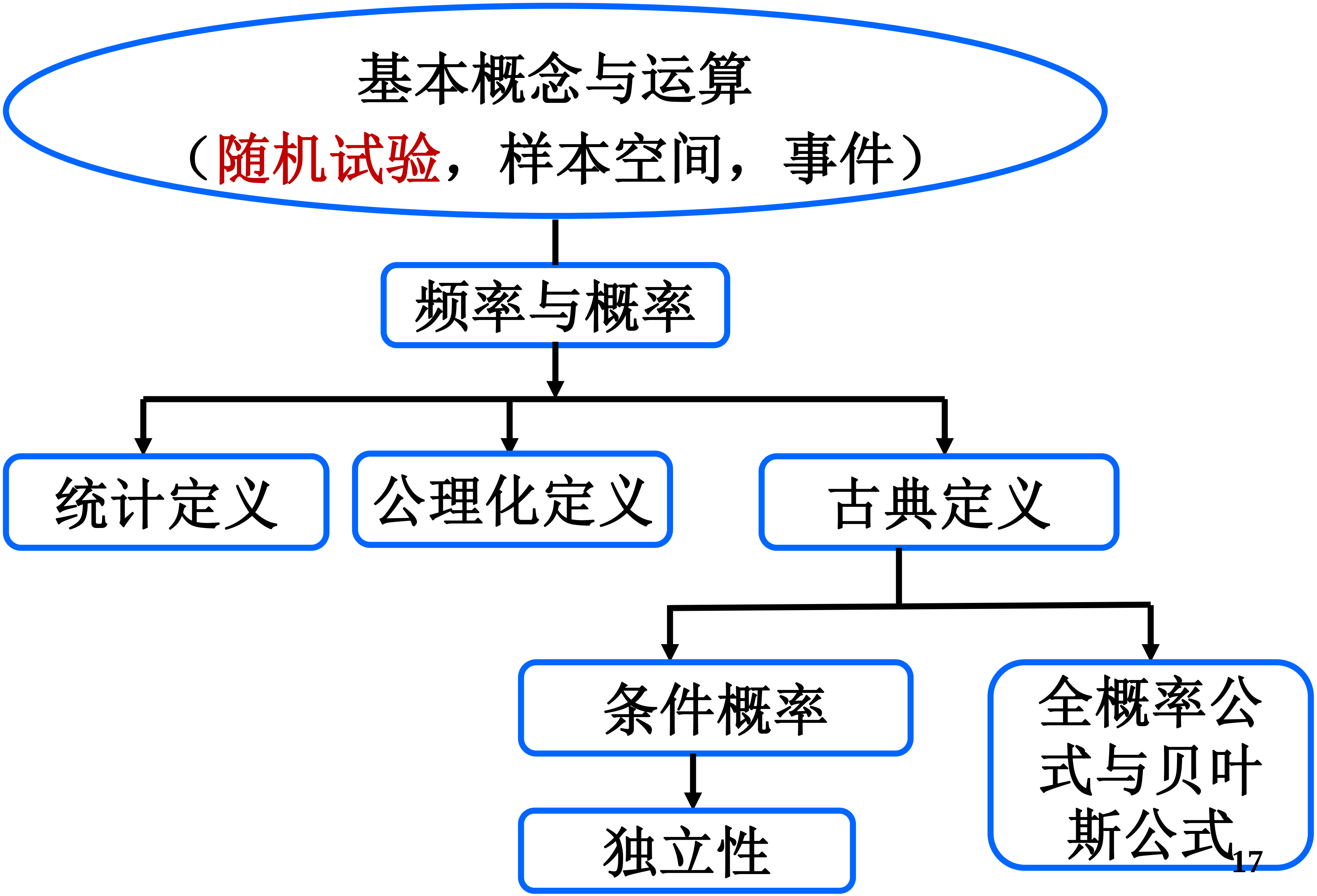
主要内容

- 课程介绍
- 课程考核及安排
- 第一章 概率论的基本概念

第一章 知识结构图


基本概念与运算
（随机试验，样本空间，事件）
频率与概率
统计定义
公理化定义
古典定义
条件概率
独立性
全概率公式与贝叶斯公式

第一节 随机试验


基本概念与运算
（随机试验，样本空间，事件）
频率与概率
统计定义
公理化定义
古典定义
条件概率
独立性
全概率公式与贝叶斯公式

第一节 随机试验

为了研究随机现象，就要对客观事物进行**试验或观察**（对某一事物某一特征的观察也认为是一种试验），记为E。

例如：

E1：掷一枚硬币观察正面、反面出现的情况。

E2：记录一小时内，到某保险公司投保的户数。

E3：射手射击一个目标，直到射中为止，观察其射击的次数。

E4：从一批产品中抽取十件，观察其次品数。

E5：抛一颗骰子，观察其出现的点数。

第一节 随机试验

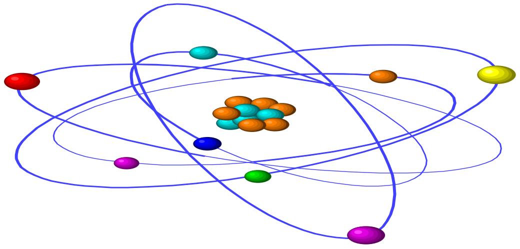
具有以下三个特性的试验称为随机试验：

① 可以在相同的条件下重复进行；

② 每次试验的结果具有多种可能性，并且能事先明确试验的所有可能结果；

③ 在每次试验之前不能确定哪一个结果会出现。

通过研究随机试验来研究随机现象。



备注：后面所提到的试验都是指随机试验。

第二节 样本空间、随机事件

基本概念与运算

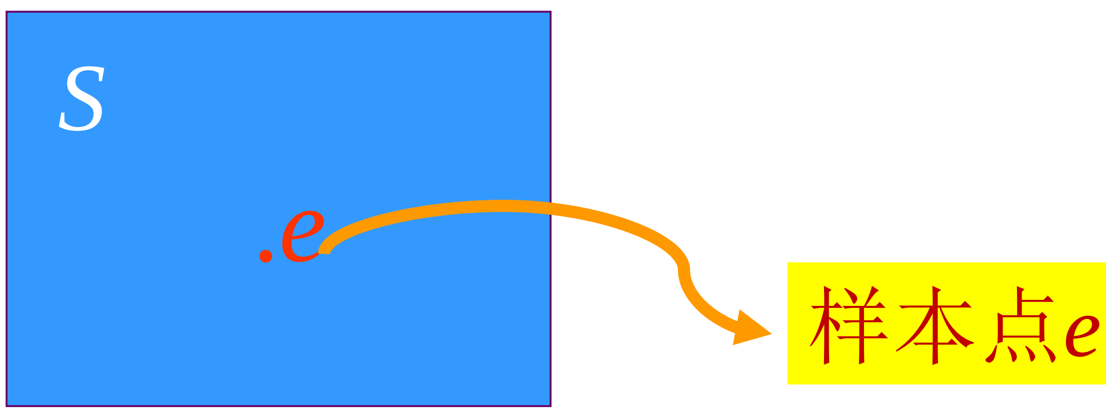
（随机试验，样本空间，事件）

频率与概率

- 统计定义
- 公理化定义
- 古典定义
  - 条件概率
    - 独立性
  - 全概率公式与贝叶斯公式

样本空间

- **样本空间**：随机试验E的所有可能结果组成的集合称为E的样本空间，记为S。
- **样本点**：样本空间的元素，即E的每一个结果称为样本点。



样本空间

例1： 写出第一节中例 E1～ E5的样本空间S。

$$S_1 = \{正面, 反面\}$$

$$S_2 = \{0, 1, 2, 3, \cdots\cdots\}$$

$$S_3 = \{1, 2, 3, \cdots\cdots\}$$

$$S_4 = \{0, 1, 2, 3, \cdots\cdots, 10\}$$

$$S_5 = \{1,2, 3,4, 5, 6\}$$

E1：掷一枚硬币观察正面、反面出现的情况。
E2：记录一小时内，到某保险公司投保的户数。
E3：射手射击一个目标，直到射中为止，观察其射击的次数。
E4：从一批产品中抽取十件，观察其次品数。
E5：抛一颗骰子，观察其出现的点数。

备注：样本空间元素是由试验目的所确定的，不同的试验目的其样本空间也是不一样的。

样本空间

例2：若试验是将一枚硬币**抛掷两次**，试写出该试验的样本空间。（硬币正面记为H，反面记为T）

解：

样本空间

北京邮电大学
Beijing University of Posts and Telecommunications

例2：若试验是将一枚硬币抛掷两次，试写出该试验的样本空间。（硬币正面记为H，反面记为T）

解：

$$S=\{(H,H),\ (H,T),\ (T,H),\ (T,T)\}$$

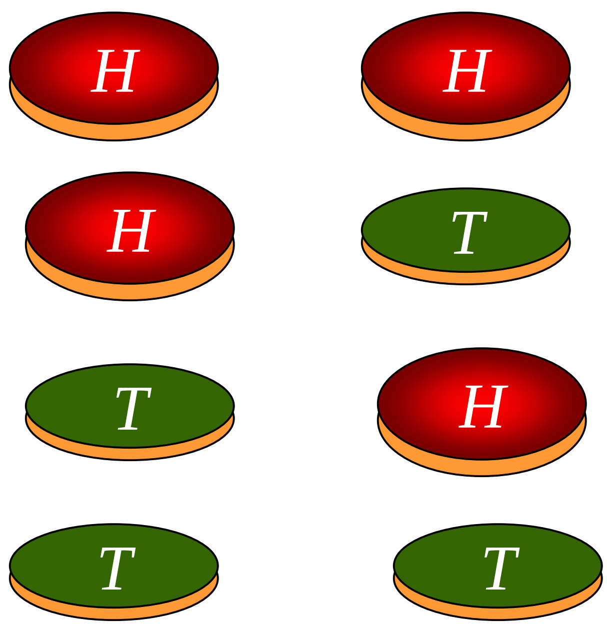
第1次　　第2次

(H,H):

(H,T) :

(T,H) :

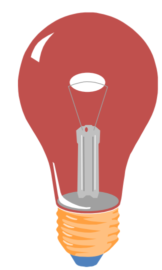
(T,T) :



样本空间

例3：若试验是测试某灯泡的寿命（以h计），试写出该试验的样本空间。

解：因为该试验的样本点是一非负数，又由于不能确知寿命的上界，所以可以认为任一非负实数都是一个可能结果，故得样本空间为：

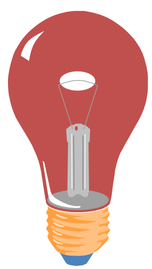


样本空间

例3：若试验是测试某灯泡的寿命（以h计）,试写出该试验的样本空间。

解：因为该试验的样本点是一非负数，又由于不能确知寿命的上界，所以可以认为任一非负实数都是一个可能结果，故得样本空间为：

$$S = \{t \mid t \geq 0\}$$



样本空间

例4：若试验是记录某地一昼夜的最高温度和最低温度, 试写出该试验的样本空间。

解：用x表示最低温度（以℃计），y表示最高温度（以℃计），设这一地区的温度不会小于T_0，不会大于T_1，故得样本空间为：

样本空间

例4：若试验是记录某地一昼夜的最高温度和最低温度，试写出该试验的样本空间。

解：用x表示最低温度（以℃计），y表示最高温度（以℃计），设这一地区的温度不会小于T_0，不会大于T_1，故得样本空间为：

$$S=\{(x,y)\mid T_0\le x\le y\le T_1\}\text{。}$$

随机事件

- 随机事件：称试验E的样本空间S的子集为E的随机事件，简称事件。

例：若试验E是测试某灯泡的寿命（以h计），E的样本空间：$S = \{t \mid t \geq 0\}$。

若规定灯泡寿命小于500的为次品，则在E中关心的灯泡寿命是否有$t \geq 500$，满足这一条件的样本点组成S的一个子集：$A=\{t \mid t \geq 500\}$。则称A为E的一个随机事件。

随机事件

- **随机事件**：称试验E的样本空间S的子集为E的随机事件，简称事件。
- **事件发生**：在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一事件发生。

例如，在掷一颗骰子观察点数的试验中，事件A={2, 4}。若在一次试验中，“掷出4点”，则称A发生。

随机事件

- 基本事件：由一个样本点组成的单点集称为基本事件。

例如，抛一枚硬币的试验中，有两个基本事件{H}和{T}。

总结：

- 样本空间就是全体基本事件的集合；
- 随机事件是某些基本事件的集合，它是样本空间的子集。

随机事件

例题：请用样本点的集合表示出下列各事件。

（1）10 件产品中仅标号为10的产品为废品，从中任取两件产品，有一件废品。

（2）在一批灯炮中任取一只，其寿命不大于100小时。

随机事件

例题：请用样本点的集合表示出下列各事件。

（1）10 件产品中仅标号为10的产品为废品，从中任取两件产品，有一件废品。

$$A=\{\text{任取两件产品中有一件是废品}\}$$
$$=\{(1,10),(2,10),\cdots\cdots,(9,10)\}$$

（2）在一批灯炮中任取一只，其寿命不大于100小时。

$$B=\{t \mid 0 \le t \le 100\ \text{小时}\}$$

Beijing University of Posts and Telecommunications

随机事件

- **必然事件：在每次随机试验中一定会出现的事件称之为必然事件。它是特殊的随机事件，包含了全部的基本事件，即为样本空间 S。**

例：E：从一批产品中取出十件，观察其次品数。
显然，S＝{0,1,2,……,10}，A={次品数小于12}是一个必然事件，它就是 S。

- **不可能事件：在任何试验中都不会出现的事件称为不可能事件。**它是特殊的随机事件，它**不包含任何基本事件，实际上它是空集∅。**

例：在以上例子中，B＝ {次品数大于15)}就是一个不可能事件，即 B是空集。

归纳

例： 在掷骰子试验中，观察掷出的点数 。

事件

基本事件

（相对于观察目的**不可再分解**的事件 ）

事件 A_i ={掷出 i 点}

i =1, 2, 3 ,4, 5, 6

复合事件

（两个或一些基本事件并在一起所构成的事件）

事件 B={掷出奇数点}

随机事件间的关系及其运算

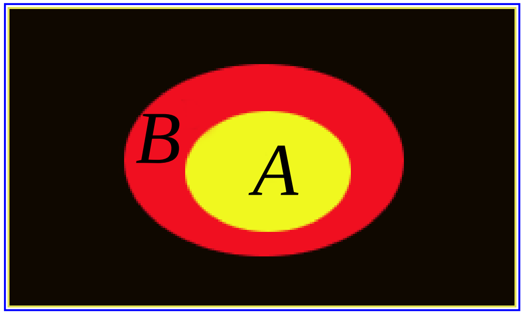
- 设试验 E 的样本空间为 S，而A、B 是 S 的子集。
- 事件的包含：如果事件A发生必然导致事件B发生（A中的每个样本点都包含在 B 中），则称 事件B包含事件A或A含于事件B。记作：$B \supset A$ 或 $A \subset B$

注：

▲ $B \supset A$ 的一个等价说法:

如果B不发生必然导致A也不发生。

▲显然对任意事件A有 $\Phi \subset A \subset S$

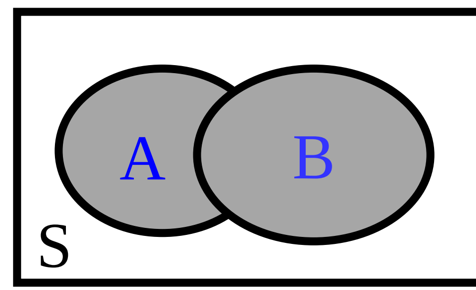


$A \subset B$

随机事件间的关系及其运算

- **事件的相等：** 若事件A, B满足 A ⊂ B且B ⊂ A，则称**事件A与B相等**，记作 A=B（A与B包含的样本点完全相同）。

- **事件的和（并）：** 事件 $A+B$或$A\cup B=\{x|x\in A$或$x\in B\}$称为事件A与B的和事件。当且仅当A，B中至少有一个发生时，事件A∪B发生。



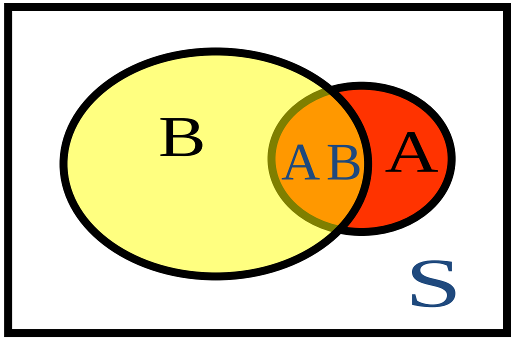
注：

▲ 称 $\bigcup_{k=1}^{n} A_k$为n个事件$A_1,\cdots,A_n$的和事件

▲ 称 $\bigcup_{k=1}^{\infty} A_k$为可列个事件$A_1,A_2,\cdots$的和事件

随机事件间的关系及其运算

- **事件的积（交）**：事件 $A \cap B = \{x | x \in A \text{且} x \in B\}$ 称为事件**A与B的积事件，当且仅当A，B同时发生时，事件A∩B发生，A∩B也记作AB。**



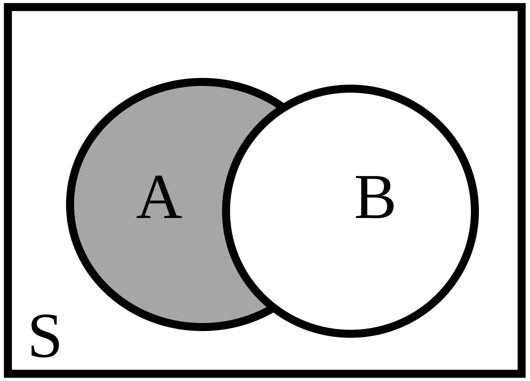
AB

注： ▲ 称 $\bigcap_{k=1}^{n} A_k$ 为n个事件$A_1, \cdots, A_n$的积事件

▲ 称 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件$A_1, A_2, \cdots$的积事件

随机事件间的关系及其运算

- **事件的差：** 事件 $A-B=\{x \mid x\in A 且 x\notin B\}$ 称为事件A与事件B的差事件。当且仅当A发生，B不发生时，事件A-B发生。

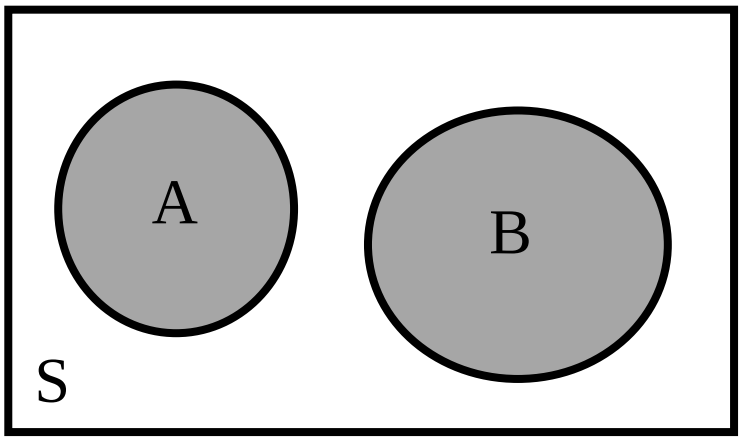


$A-B$

注： 它是由属于A但不属于B的那些样本点构成的集合。

随机事件间的关系及其运算

- 互不相容（互斥）事件：若$AB=\Phi$，则称事件A与B是互不相容的，或互斥的。这指的是事件A与事件B不能同时发生。基本事件是两两互不相容的。

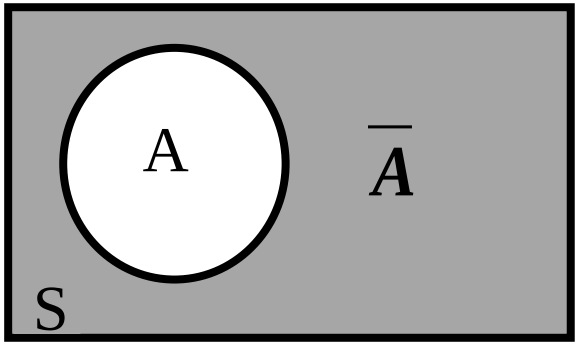


$$AB=\Phi$$

随机事件间的关系及其运算

- 对立事件（逆事件）：若 $A\cup B=S$ 且 $A\cap B=\Phi$，则称事件A与事件B互为逆事件，又称事件A与事件B互为对立事件。即，每次试验，事件A，B中必有一个发生，且仅有一个发生。

- A的对立事件记为：$\overline{A}=S-A$



$\overline{A}=S-A$

- 显然：$A\overline{A}=\Phi, A+\overline{A}=S, \overline{A}=S-A, \overline{\overline{A}}=A$

随机事件间的关系及其运算

★ 事件运算所满足的下述定律:

1. 交换律: $A \cup B = B \cup A, A \cap B = B \cap A$

2. 结合律:

$$A \cup (B \cup C) = (A \cup B) \cup C$$

$$A \cap (B \cap C) = (A \cap B) \cap C$$

3. 分配律:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

4. 德摩根律:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

随机事件间的关系及其运算

例：随机试验E：对某一目标接连进行两次射击, 记

$$A_i = \{ 第i次射击命中目标 \}，\ i = 1,2$$

试用事件间的关系和运算表示下列各事件：

（1）第 i 次射击未命中目标 $i = 1,2$

（2）$B_j = \{ 两次射击恰好有j次命中目标 \}$，$j = 0,1,2$

（3）$C_k = \{ 两次射击至少有k次命中目标 \}$，$k = 0,1,2$

解：

随机事件间的关系及其运算

例：随机试验E：对某一目标接连进行两次射击, 记

$$A_i=\{\text{第}i\text{次射击命中目标}\},\quad i=1,2$$

试用事件间的关系和运算表示下列各事件：

（1）第 i 次射击未命中目标 $i=1,2$

（2）$B_j=\{\text{两次射击恰好有}j\text{次命中目标}\},\quad j=0,1,2$

（3）$C_k=\{\text{两次射击至少有}k\text{次命中目标}\},\quad k=0,1,2$

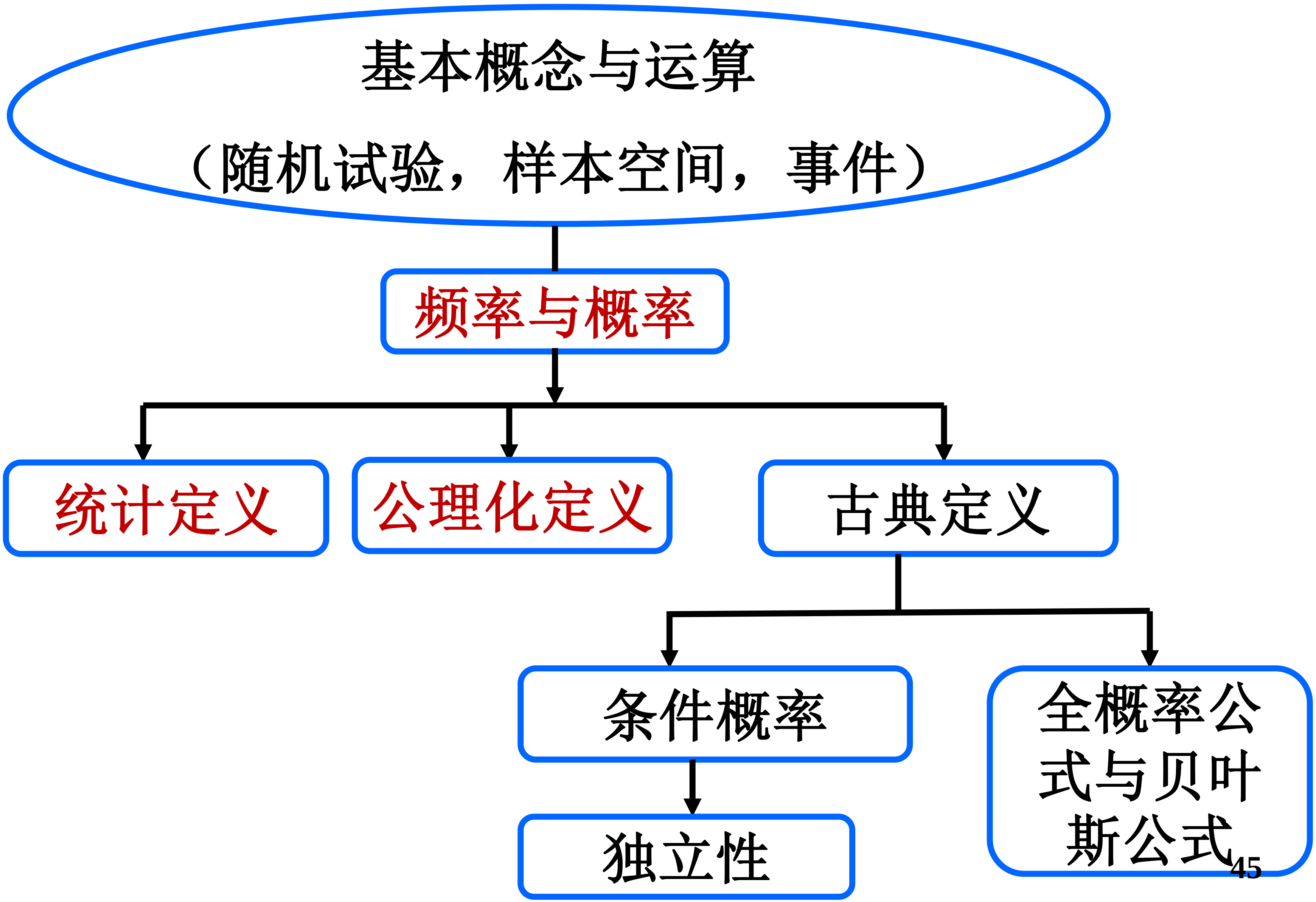
解：（1） 依题意：$\overline{A_i}=\{\text{第}i\text{次射击未命中目标}\}$ ，$i=1,2$

（2） $B_0=\overline{A_1}\,\overline{A_2}$，$B_1=A_1\overline{A_2}\cup\overline{A_1}A_2$，$B_2=A_1A_2$

（3）$C_0=B_0\cup B_1\cup B_2=S$

$C_1=B_1\cup B_2$，　$C_2=B_2=A_1A_2$

第三节 频率与概率


基本概念与运算
（随机试验，样本空间，事件）
频率与概率
统计定义
公理化定义
古典定义
条件概率
独立性
全概率公式与贝叶斯公式

频率

1．频率的定义：

在 n 次试验中,事件A发生的次数 n_A 称为事件A的频数,而比值 $\frac{n_A}{n}$ 称为事件A发生的频率,记作：

$$f_n(A)=\frac{n_A}{n}$$

2．频率的性质：

(1) $0\le f_n(A)\le 1$(非负性)　(2) $f_n(S)=1$(规范性)

(3) 若 $A_1,A_2,\cdots A_K$ 是两两互不相容的事件，则

$$f_n(A_1\cup A_2\cup\cdots A_k)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k)$$

（有限可加性）

频率

3．频率的稳定性

概率统计定义

在不变的条件下,重复进行 n 次试验,事件A发生的频率 n_A/n 稳定地在某一常数 p 附近摆动；并且 n 越大,逐渐稳定于常数p。则称常数 p为**事件A 在该条件下发生的概率**。(简称:频率的稳定值为该事件的概率)

记作：**P(A)**。

注：这种“频率稳定性”即通常所说的统计规律性。

概率

公理化定义

1. 概率的定义：

设E是随机试验，S是它的样本空间。对于E的每一个事件A，赋予一个实数，记为 $P(A)$，称为事件A的概率，如果它满足以下三个条件：

（1）对于每一事件A有: $0 \le P(A) \le 1$ 非负性

（2）对于必然事件S，有$P(S) = 1$ 规范性

（3）设$A_1,A_2,\dots$是两两互不相容的事件，

则$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$ 可列可加性

2． 概率的性质

性质1 $\boldsymbol{p(\Phi)=0}$

[证]：令 $A_1=A_2=\cdots=\Phi$，则 $\boldsymbol{A_iA_j}=\Phi$

$$\boldsymbol{P(\Phi)=P(A_1\cup A_2\cup A_3\cup\cdots)=\sum_{i=1}^{\infty}P(A_i)=\sum_{i=1}^{\infty}P(\Phi)}$$

由于 $\boldsymbol{P}(\Phi)\geq 0$ 故，$\boldsymbol{P}(\Phi)=0$

概率

性质2 (有限可加性)　若 $A_1, A_2, \cdots A_n$ 是两两互不相容事件, 则有:

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

[证]:　由概率定义中的可列可加性。

令: $A_{n+1} = A_{n+2} = \cdots = \Phi$ ，则有:

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(\bigcup_{k=1}^{\infty} A_k)$$

$$= \sum_{k=1}^{n} P(A_k) + \sum_{k=n+1}^{\infty} P(A_k)$$

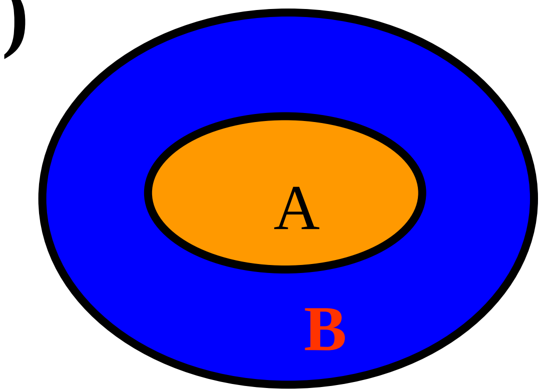
$$= \sum_{k=1}^{n} P(A_k) + 0 = \sum_{k=1}^{n} P(A_k)$$

概率

性质3：若 $A \subset B$,则有

(可减性) $P(B-A) = P(B) - P(A)$

(单调性) $P(B) \geq P(A)$



[证]：(1) $\because B = A \cup (B-A)$

并且A与B-A是互不相容的，即：

$$A(B-A) = \Phi$$

$\therefore$ 由性质 2 得：

$$P(B) = P(A) + P(B-A),$$

$$P(B-A) = P(B) - P(A)$$

概率

性质3：若 $A \subset B$,则有

(可减性) $P(B-A)=P(B)-P(A)$

(单调性) $P(B) \geq P(A)$

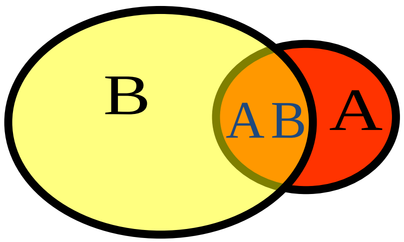
[证]：(2) 由概率定义可知：$P(B-A) \geq 0$

$\therefore$由(1)得 $P(B)-P(A) \geq 0$

即：$P(B) \geq P(A)$

性质4 (加法定理) 设 A, B为任意两个事件, 则有:

$$P(A\cup B)=P(A)+P(B)-P(AB)$$

[证]:由图 

可知: $A\cup B=A\cup(B-AB)$

并且: A与$B-AB$是互不相容的

∴由性质2与性质3得:

$$P(A\cup B)=P(A)+P(B-AB)$$

$$=P(A)+P(B)-P(AB)$$

概率

注：性质4可推广到多个事件:

$$P(A_1 \cup A_2 \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i \le j \le n} P(A_i A_j)$$

$$+ \sum_{1 \le i < j < k \le n} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n)$$

比如:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$$

$$- P(A_1 A_2) - P(A_1 A_3) - P(A_2 A_3) + P(A_1 A_2 A_3)$$

概率

性质5 对任意事件A有：$P(\overline{A})=1-P(A)$

[证]：$\because A\cup\overline{A}=S$ 并且 $A\overline{A}=\Phi$

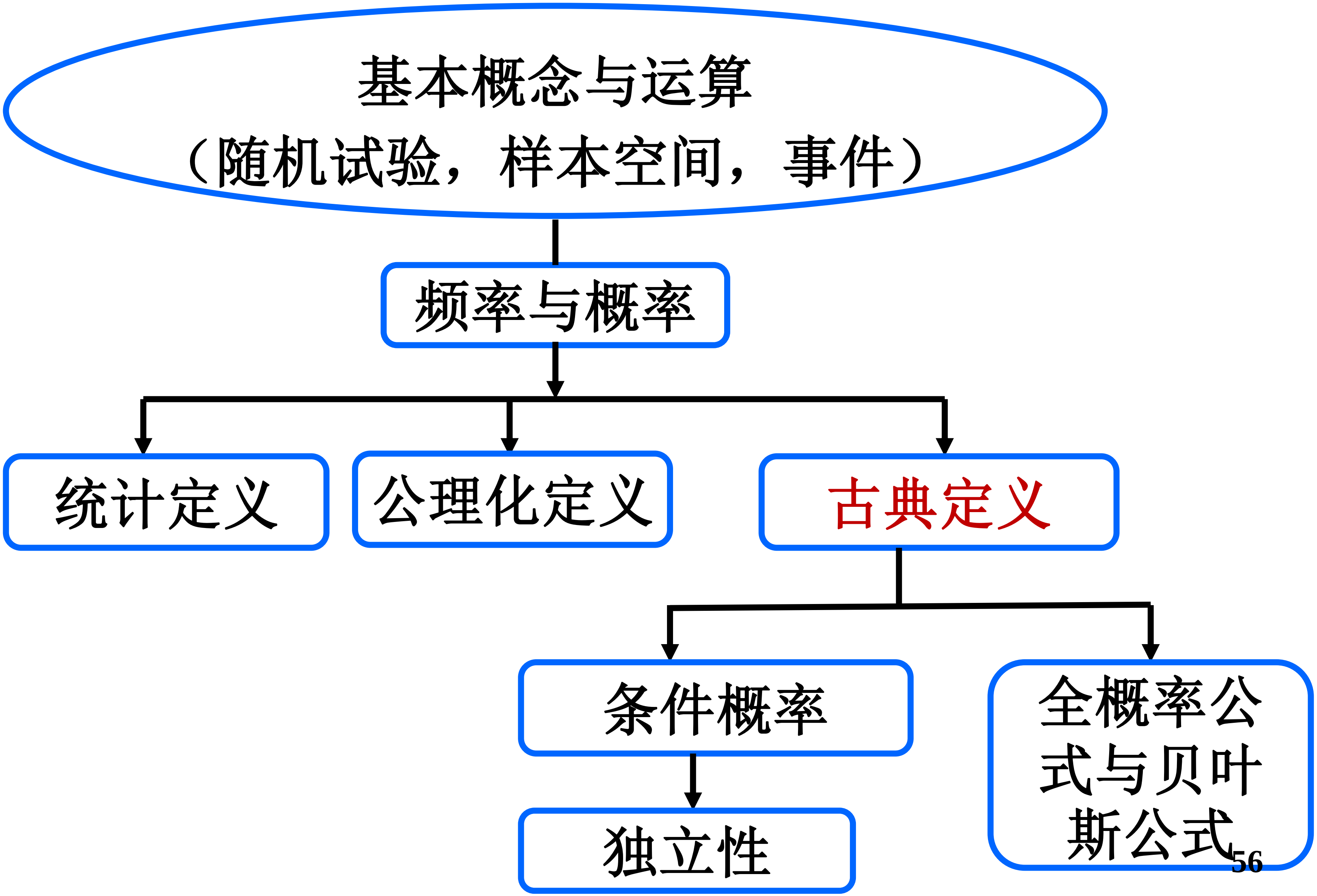
$\therefore$ 由性质2（有限可加性）得：

$$P(S)=P(A\cup\overline{A})=P(A)+P(\overline{A})$$

又 $\because$ $P(S)=1$ 即：$P(\overline{A})=1-P(A)$

注：性质5在概率的计算上很有用，如果正面计算事件A的概率不容易，而计算其对立事件的概率较易时，可以先计算$P(\overline{A})$，再计算$P(A)$。

第四节 等可能概型（古典概型）



等可能概型试验（古典概型）

一般，如果随机试验 E 具有：

(1) 有限性：它的样本空间的元素只有有限个；

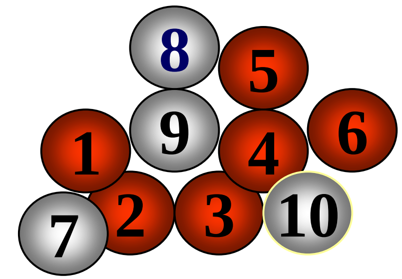
(2) 等可能性：在每次试验中，每个基本事件发生的可能性相同。

则称随机试验E为等可能概型。它在概率论发展初期曾是主要的研究对象，所以也称为古典概型。

等可能概型试验（古典概型）

示例：一个箱子里有10个球，其中白球4个，红球有6个。随机地从箱中取一个球。

E1:观察取出的球的号码。



$A_i=\{\text{取出球的号码为}i\}$，$i=1,2,\cdots 10$

$P(A_i)=\dfrac{1}{10}, i=1,2,\cdots 10$。

等可能概型

E2:观察取出红球、白球的情况。

$B=\{\text{取出球为红球}\}$，$\bar{B}=\{\text{取出球为白球}\}$

$P(B)=\dfrac{6}{10}=\dfrac{3}{5}$，$P(\bar{B})=\dfrac{2}{5}$。

非等可能概型

古典概型中事件概率的计算公式

定义：设E是古典随机试验，S是它的样本空间，

$S=\{e_1,e_2,\cdots,e_n\}$，若事件A包含 k 个基本事件

即$A=\{e_{i_1}\}\cup\{e_{i_2}\}\cup\cdots\cup\{e_{i_k}\}$ $(1\le i_1<i_2<\cdots i_k\le n)$

则称 $P(A)=\sum_{j=1}^{k}P(\{e_{i_j}\})=\frac{k}{n}=\frac{A\text{包含的基本事件}}{S\text{中基本事件}}$

为事件 A 的概率。

概率的古典定义

注：▲教材 P10 给出了它的推导过程。

古典概型中事件概率的计算公式

例：一次掷两颗骰子，观察点数和小于5的情况。

事件B ={ 两颗骰子点数之和小于 5 }

i. 若样本空间 S={2, 3, …, 12}，则 $B=\{2,3,4\}$

从而，P(B)=3/11

错误的，因为不是古典概型，S中每个基本事件发生的可能性不同

ii. 若样本空间 $S=\{(i,j) \mid i, j = 1, 2, \ldots, 6\}$，则

B={(1,1), (1,2), (1,3), (2,1), (2,2) , (3,1) }

从而，P(B)=6/36

▲ 排列组合是必备的基础知识。

关于排列组合知识复习——一、基本计数原理

1. 加法原理

设完成一件事有m种方式，

第一种方式有n_1种方法，

第二种方式有n_2种方法，

…，

第m种方式有n_m种方法，

则完成这件事总共有$n_1+n_2+\dots+n_m$种方法。

无论通过哪种方式的方法都可以完成这件事。

关于排列组合知识复习——一、基本计数原理

2. 乘法原理

设完成一件事有m个步骤，

第一个步骤有$n1$种方法，

第二个步骤有$n2$种方法，

…，

第m个步骤有n_m种方法，

则完成这件事共有

$$n_1 \times n_2 \times \cdots \times n_m$$

种不同的方法（组合）。

必须通过每一步骤，才算完成这件事。

关于排列组合知识复习——一、基本计数原理

加法原理和乘法原理是两个很重要计数原理，它们不但可以直接解决不少具体问题，也是推导下面常用排列组合公式的基础。同时它们也是计算古典概率的基础（给出所有具备等概率特征的基本事件）。

关于排列组合知识复习——二、排列、组合的几个简单公式

1、排列：从n个不同元素取k个 $(1 \le k \le n)$

的不同排列总数为：

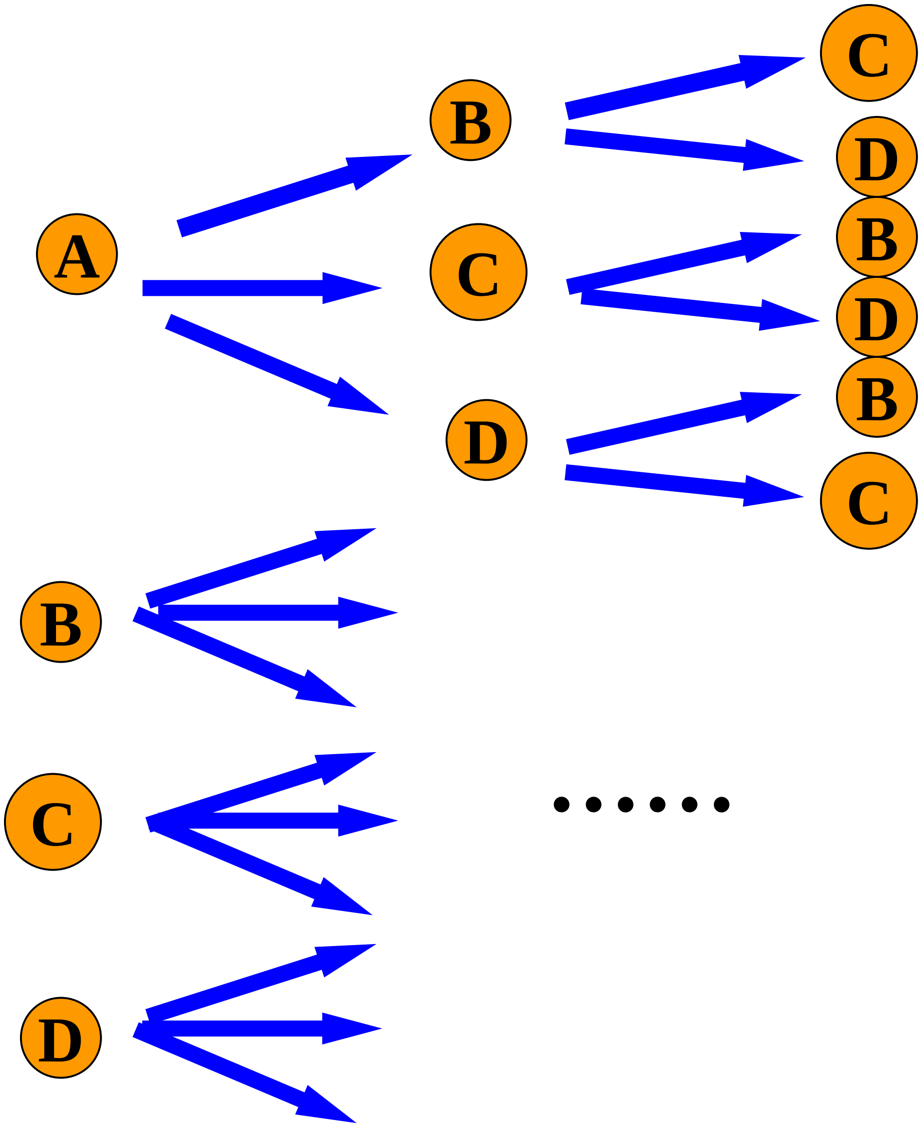
$$P_n^k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

$k = n$ 时称 全排列：

$$P_n^n = P_n = n(n-1)(n-2)\cdots 2 \cdot 1 = n!$$

示例 1—不放回抽样

第1次选取 第2次选取 第3次选取



例如：n=4, k =3

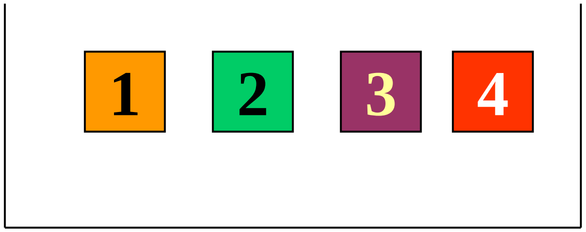
则选排列、全排列为：

$$P_4^3 = 4 \cdot 3 \cdot 2 = 24$$

$$P_4 = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

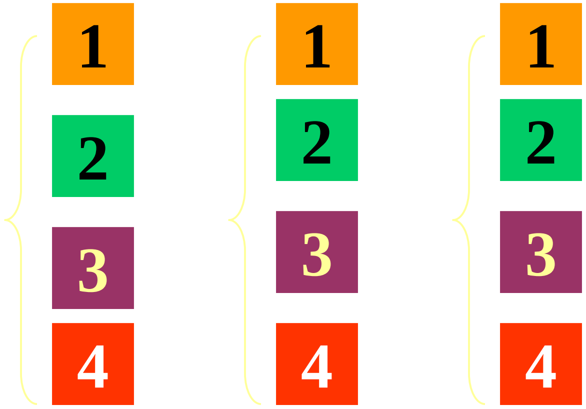
示例 2—放回抽样

从装有4张卡片的盒中有放回地摸取3张



n=4,k =3

第1张 第2张 第3张



共有4·4·4=4^3种可能取法

从n个不同元素取 k个（允许重复）的不同排列总数为：

$$n \cdot n \cdots n = n^k$$

可重复排列

关于排列组合知识复习——二、排列、组合的几个简单公式

2、组合: 从n个不同元素取 k个

$(1\le k\le n)$的不同组合总数为:

C_n^k 常记作 $\binom{n}{k}$，称为组合系数。

$$P_n^k = C_n^k \cdot k!$$

排列与组合的关系:
排列总数=组合总数*每种组合下的排列数

$$C_n^k = \frac{P_n^k}{k!} = \frac{n!}{(n-k)!k!}$$

组合不再考虑k个元素的顺序

例 1. 设有30件产品,其中有4件是次品,现从中任取3件

求：(1) 恰有 2 件次品的概率

(2) 至少有一件次品的概率

解：(1)设A：任取3件恰有两件是次品

n：无论取出的是否是次品,它总是从30件产品中取3件,故它的样本空间总数是: C_{30}^{3}

k：恰有2件次品应从次品中取, 剩下的另一件正品应从26件正品中取到。故有: $C_4^2 C_{26}^1$

从而：$$P(A)=\frac{k}{n}=\frac{C_4^2\cdot C_{26}^1}{C_{30}^3}=\frac{156}{4060}\approx 0.038$$

(2) 设 B：任取3件至少有一件次品

$n:$ C_{30}^{3}

$k:$ 次品至少有一件指的是1件，2件，3件，三种情况，故有：

$$C_4^1 \cdot C_{26}^2 + C_4^2 \cdot C_{26}^1 + C_4^3 \cdot C_{26}^0$$

从而：
$$P(B) = \frac{k}{n} = \frac{C_4^1 \cdot C_{26}^2 + C_4^2 \cdot C_{26}^1 + C_4^3 \cdot C_{26}^0}{C_{30}^3}$$

$$= \frac{1460}{4060} \approx 0.36$$

关于排列组合知识复习——二、排列、组合的几个简单公式

3、超几何分布的概率公式:

在N件产品中不放回的抽取n件，所有可能的取法共有C_N^n，每一种取法为一基本事件，且每个基本事件发生的可能性相同。

又因在D件次品中取k件，所有可能的取法有C_D^k种，在$N-D$件正品中取$n-k$件所有可能的取法有C_{N-D}^{n-k}种。

由乘法原理，在N件产品中不放回的抽取n件，其中有k件次品的概率为$\frac{C_D^k C_{N-D}^{n-k}}{C_N^n}$（超几何分布的概率公式）。

例 2. 盒中有6张面值相同的债券,其中有两张中奖债券,现从中任取两次,每次取一张,考虑两种取法:

(1). 有放回地取：第一次取出观察后放回盒中混合均匀后再取第二次 (放回抽样)

(2). 无放回地取：第一次取出后不放回盒中, 第二次从剩余的债券中再取一张 (不放回抽样)

求：分别就两种抽样方式求取到的两张都是中奖的债券的概率？

解：显然，本题属古典概型。

(1). 有放回地抽取：设A：取到的两张都是中奖券

n：第一次从盒中取，不论是否是中奖券，总是从 6 张中取一张，第二次再从盒中取，仍是有 6 张券可供抽取，故有：

$$P_6^1 \cdot P_6^1 = 36\,(种)$$

k：中奖券有 2 张，第一次取有 2 张可供抽取,第二次取仍有 2 张可供抽取，故有：$P_2^1 \cdot P_2^1 = 4\,(种)$

从而：$$P(A) = \frac{k}{n} = \frac{4}{36} = \frac{1}{9} = 0.111$$

（2). 不放回地抽取：

$n: \ P_6^1 \cdot P_5^1 = 30$ (种), $k: \ P_2^1 \cdot P_1^1 = 2$ (种),

从而：$$P(A) = \frac{k}{n} = \frac{2}{30} = \frac{1}{15} = 0.067$$

注:▲在此例中若将取法改为“一次抽取两张”，其它条件不变则有：$$P(A) = \frac{C_2^2}{C_6^2} = \frac{1}{15} = 0.067$$

▲“不放回地抽取两次, 每次取一张” 相当于 “一次抽取两张”。故在许多问题中如果不是有放回地抽样, 就统称为“任意取出”多少个。

例3 把C、C、E、E、I、N、S七个字母分别写在七张同样的卡片上，并且将卡片放入同一盒中，现从盒中任意一张一张地将卡片取出，并将其按取到的顺序排成一列，假设排列结果恰好拼成如下英文单词的概率：

S C I E N C E

解: 设A：排列结果恰好拼成英文单词 *S C I E N C E*

n：七个字母的排列总数为 7!

k：拼成英文单词*SCIENCE* 的情况数为：

$2\times2=4$　　故该结果出现的概率为：

$$P(A)=\frac{4}{7!}=\frac{1}{1260}\approx0.00079$$

这个概率很小，这里算出的概率有如下的实际意义：

如果多次重复这一抽卡试验，则我们所关心的事件在1260 次试验中大约出现 1 次。

例4　某接待站在某一周曾接待过12次来访，已知这12次接待都是在周二和周四进行的。

问：是否可以推断接待时间是有规定的？

解：假设接待站的接待时间没有规定，而各来访者在一周的任一天去接待站是等可能的。

设 A：12次接待来访者都在周二和周四

n：一周共有7天，12次来访者均可在七天中任一天都可去接待站，相当于从七天中要接待12次且可以 重复日期，故有：7^{12} (种)

k : 12次接待来访者只能放在周二和周四两天之中,

故有： 2^{12}

从而 $P(A)=\dfrac{k}{n}=\dfrac{2^{12}}{7^{12}}=0.0000003$ (千万分之三)

注: 由“实际推断原理”：概率很小的事件在一次试验中实际上几乎是不发生的。

可推断到：“假设接待站接待时间是没有规定”这一说法是不正确的，因为千万分之三的小概率事件竟然在假设条件下发生了。

故得出：接待站的接待时间是有规定的，而不是每天都接待的。

例 5　有 r 个人，设每个人的生日是365天的任何一天是等可能的，试求事件“至少有两人同生日”的概率。

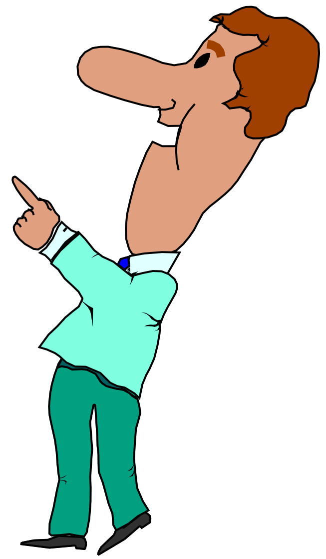
解：令　A={至少有两人同生日}

则　$\overline{A}$={ r 个人的生日都不同}

为求$P(A)$，先求 $P(\overline{A})$：

$$P(\overline{A}) = \frac{P_{365}^{r}}{(365)^{r}}$$

则有：$P(A) = 1 - P(\overline{A}) = 1 - \dfrac{P_{365}^{r}}{(365)^{r}}$



统计表

人数	至少有两人同生日的概率
20	0.411
21	0.444
22	0.476
23	0.507
24	0.538
30	0.706
40	0.891
50	0.970
60	0.994

所有这些概率都是在假定一个人的生日在 365天的任何一天是等可能的前提下计算出来的。

根据统计表结果，当人数超过23时，打赌说至少有两人同生日是有利的。

生日问题

例6　n双相异的鞋共$2n$只，随机地分成n堆，每堆2只。问:“各堆都自成一双鞋”(事件A)的概率是多少？

解:

例6 n双相异的鞋共$2n$只，随机地分成n堆，每堆2只。问:“各堆都自成一双鞋” (事件A)的概率是多少？

解： 把$2n$只鞋分成n堆，每堆2只的分法总数为：

$$\frac{(2n)!}{2!2!\cdots 2!} = \frac{(2n)!}{2^n}$$

1、考虑排列总数；
2、划分n堆；
3、不考虑每堆2双鞋的顺序

而出现事件A的分法数为$n!$，故得：

$$P(A) = \frac{n!}{(2n)!/2^n} = \frac{n! \cdot 2^n}{(2n)!}$$

例7　甲、乙两人先后从52张牌中各抽取13张,
求甲或乙拿到4张A的概率。
1) 甲抽后不放回，乙再抽;
2) 甲抽后将牌放回，乙再抽.

解：设 A={甲拿到4张A}
B={乙拿到4张A}

所求为 $P(A\cup B)$

1) 此时A、B互斥

则：$P(A\cup B)=P(A)+P(B)$

计算$P(A)$和$P(B)$时用古典概型

$$=\frac{C_{48}^{9}}{C_{52}^{13}}+\frac{C_{48}^{13}\cdot C_{35}^{9}}{C_{52}^{13}\cdot C_{39}^{13}}=\frac{2C_{48}^{9}}{C_{52}^{13}}$$

2) 此时*A*、*B*相容

则： $P(A+B)=P(A)+P(B)-P(AB)$

$$=\frac{2C_{48}^{9}}{C_{52}^{13}}-\frac{C_{48}^{9}\cdot C_{48}^{9}}{C_{52}^{13}\cdot C_{52}^{13}}$$

例8 配对问题

某人将三封写好的信随机装入三个写好地址的信封中，问：没有一封信装对地址的概率是多少？

解：设 A_i={第 i 封信装入第 i 个信封} $i=1,2,3$

A={没有一封信装对地址}

则：$\overline{A}$={ 至少有一封信装对地址 }

直接计算 $P(A)$ 不易，则可先来计算 $P(\overline{A})$

因为：$\overline{A}=A_1+A_2+A_3$

$$\overline{A} = A_1 + A_2 + A_3$$

$$P(\overline{A}) = P(A_1 + A_2 + A_3)$$

应用加法公式

$$= P(A_1) + P(A_2) + P(A_3) - P(A_1A_2)$$

$$- P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3)$$

其中 $P(A_1) = P(A_2) = P(A_3) = \dfrac{2!}{3!} = \dfrac{1}{3}$

$$P(A_1A_2) = P(A_1A_3) = P(A_2A_3) = \frac{1}{3!} = \frac{1}{6}$$

$$P(A_1A_2A_3) = \frac{1}{3!} = \frac{1}{6}$$

代入计算 $P(\overline{A})$ 的公式中：

$$P(\overline{A})=P(A_1+A_2+A_3)$$

$$=3\cdot\frac{2!}{3!}-3\cdot\frac{1}{3!}+\frac{1}{3!}$$

$$=1-\frac{1}{2!}+\frac{1}{3!}=\frac{2}{3}$$

于是:

$$P(A)=1-P(\overline{A})$$

$$=\frac{1}{2!}-\frac{1}{3!}=\frac{1}{3}$$

推广到n封信,用类似的方法可得:把n 封信随机地装入n个写好地址的信封中, 没有一封信配对的概率为:

$$1-(1-\frac{1}{2!}+\frac{1}{3!}-\cdots+(-1)^{n-1}\frac{1}{n!})$$

$$=\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^n\frac{1}{n!}$$

例9（约会问题）两人相约7点到8点在某地会面，先到者等候另一人20分钟，过时离去。试求这两人能会面的概率。

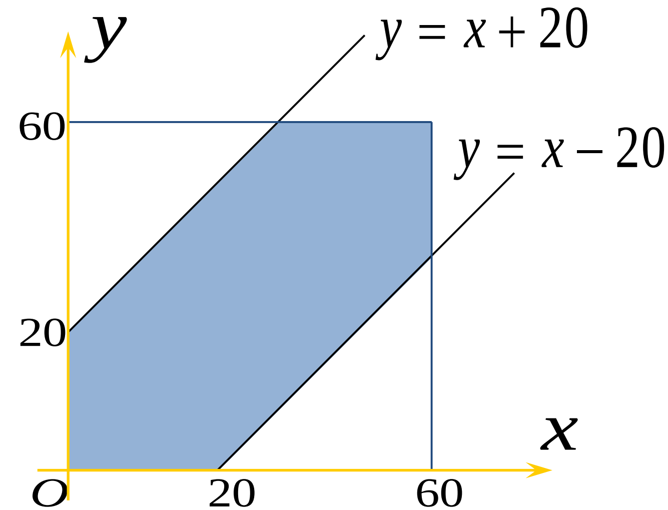
解 设 x, y 分别表示两人达到的时间，则两人能会面的充要条件是

$$|x-y| \le 20$$

$$-20 \le x-y \le 20$$

这是一个几何概型，所求概率是

$$p=\frac{60^2-40^2}{60^2}=\frac{5}{9}$$



归纳

在应用古典概型时必须注意“等可能性”的条件。

“等可能性”是一种假设，在实际应用中, 应该根据实际情况去判断是否可以认为各基本事件或样本点是等可能的。在许多场合，由对称性和均衡性，一般就可以认为基本事件是等可能的并在此基础上计算事件的概率。

第一章作业（教材第五版）：

P25：1、2、3、4
P26：8、9、10、11、13