概率论与数理统计

概率论与数理统计

概率统计是一门什么样的数学学科？

能用来研究解决哪些实际问题？

概率统计是研究解决随机问题的理论基础和重要的数学工具.

随机问题：

- **保险问题：**

某保险公司推出在校大学生意外伤害险，每位参保人投保时要交付保费500元，出险时可获得3万元赔付；已知一年中的出险率为0.15%，现有6000名新生欲参加保险.

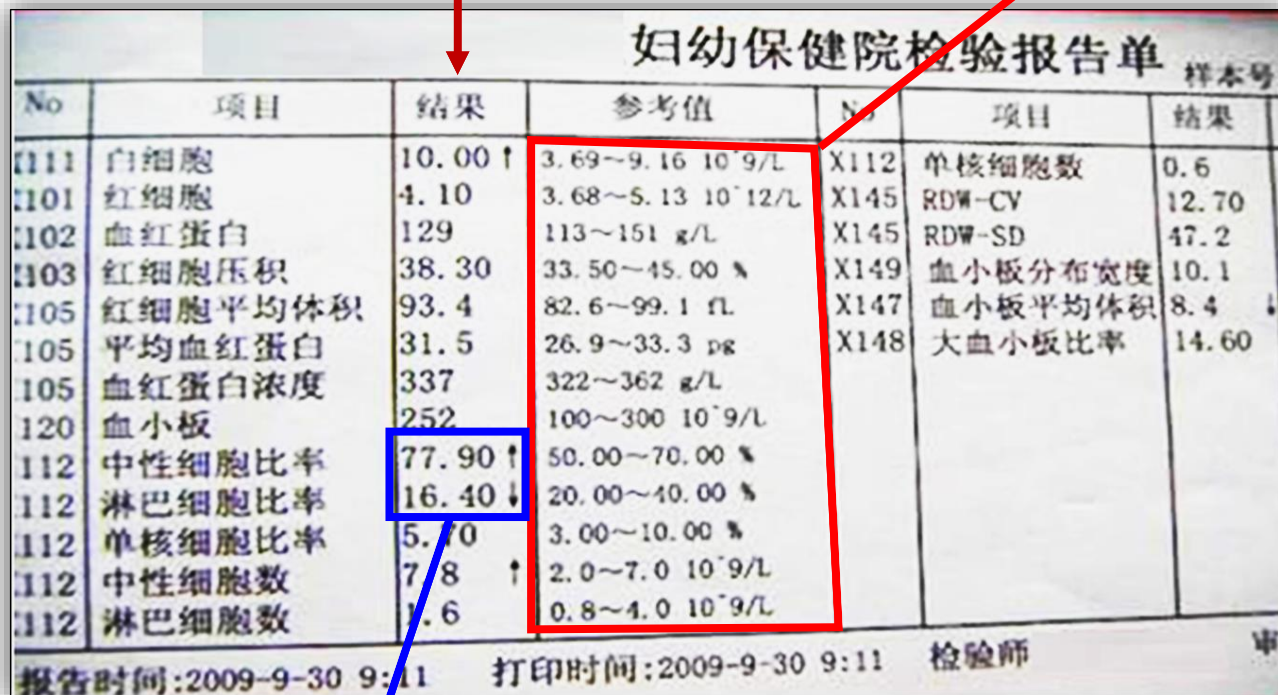
问题： 保险公司获利不少于20万元的概率是多大？

- 在校购买保险

随机问题：

- **体检报告单：**

检查结果　　　　　　　　　　　　参考值



妇幼保健院检验报告单　样本号：

No	项目	结果	参考值	No	项目	结果
111	白细胞	10.00 ↑	3.69～9.16 10^9/L	X112	单核细胞数	0.6
101	红细胞	4.10	3.68～5.13 10^12/L	X145	RDW-CV	12.70
102	血红蛋白	129	113～151 g/L	X145	RDW-SD	47.2
103	红细胞压积	38.30	33.50～45.00 %	X149	血小板分布宽度	10.1
105	红细胞平均体积	93.4	82.6～99.1 fL	X147	血小板平均体积	8.4 ↓
105	平均血红蛋白	31.5	26.9～33.3 pg	X148	大血小板比率	14.60
105	血红蛋白浓度	337	322～362 g/L			
120	血小板	252	100～300 10^9/L			
112	中性细胞比率	77.90 ↑	50.00～70.00 %			
112	淋巴细胞比率	16.40 ↓	20.00～40.00 %			
112	单核细胞比率	5.70	3.00～10.00 %			
112	中性细胞数	7.8 ↑	2.0～7.0 10^9/L			
112	淋巴细胞数	1.6	0.8～4.0 10^9/L			

报告时间:2009-9-30 9:11　打印时间:2009-9-30 9:11　检验师　审

指标异常

问题： 如何确定各项指标的参考值范围？

随机问题：

- **食堂窗口规划问题：**

学校食堂每天中午都要为全校约10000名学生提供午餐。假设每个学生在窗口打饭的时间相互独立，都服从$\lambda = 2$的指数分布。

问题： 至少需要开设多少个窗口才能保证所有学生以99%的概率在30分钟内买完饭？

概率统计是研究解决随机问题的理论基础和重要的数学工具.

确定性现象: 在一定条件下必然发生的现象.

例如: 生老病死；花开花落；上抛的石子必然落下。

随机现象: 在个别试验中其结果呈现不确定性；在大量试验中其结果又具有统计规律性的现象.

例如: 抛硬币

每次抛掷硬币之前无法确定抛掷的结果；
但大量重复抛掷, 正反面朝上的次数几乎一样。

概率统计是研究随机现象中统计规律性的一门数学学科.

概率论与数理统计的区别与联系

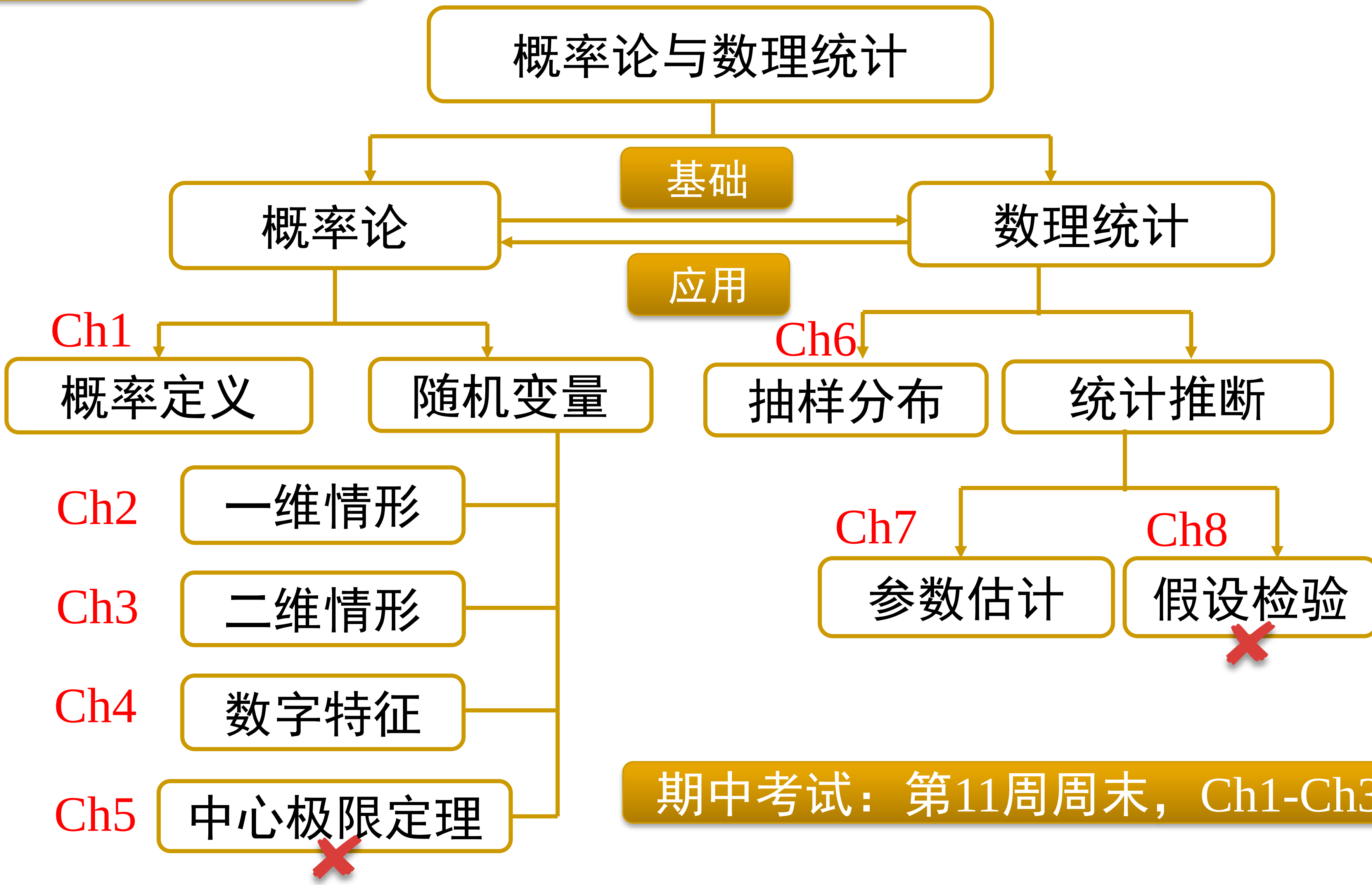
概率论与数理统计是两个完全独立的数学学科，它们有着各自不同的研究对象：

概率论主要研究的是：随机现象中事件发生的概率。

统计学主要研究的是：对实际问题中大量随机数据进行分析，整理，从而对未知量进行统计推断和预测。

概率论是统计学的基础，统计学是概率论的应用。

课程知识结构图



期中考试：第11周周末，Ch1-Ch3

第一章 随机事件与概率

教学计划：3次课-9学时

第一章 随机事件与概率

第一节 随机试验

第二节 样本空间及随机事件

- 随机试验
- 样本空间
- 随机事件
- 随机事件的关系及其运算

一. 随机试验:

试验: 为了研究随机现象，就要对客观事物进行观察, 观察的过程称为试验。记为E。 **Experiment**

例: E_1: 掷一枚硬币,观察正面H 、反面T出现的情况。

E_2: 将一枚硬币抛掷三次,观察正面H,反面T出现的情况。

E_3: 将一枚硬币抛掷三次,观察正面出现的次数

E_4: 抛一颗骰子，观察其出现的点数。

E_5: 记录120急救台一昼夜接到的呼叫次数。

E_6: 在一批灯泡中任意抽取一只，测试其寿命。

一. 随机试验:

随机试验:

具有以下三个特征的试验称为(随机)试验:

（1）试验可以在相同的条件下重复进行 (可重复性)

（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(结果的多样性)

（3）在每次试验之前不能确定哪一个结果可能会出现(结果的不确定性)

随机试验：
- （1）可重复性
- （2）结果的多样性
- （3）结果的不确定性

例： E_1： 掷一枚硬币,观察正面、反面出现的情况。

E_2： 将一枚硬币抛掷三次,观察正面H,反面T出现的情况。

E_3： 将一枚硬币抛掷三次,观察正面出现的次数

E_4： 抛一颗骰子，观察其出现的点数。

E_5： 记录120急救台一昼夜接到的呼叫次数。

E_6： 在一批灯泡中任意抽取一只，测试其寿命。

第一章 随机事件与概率

第一节 随机试验
第二节 样本空间及随机事件

- 随机试验
- 样本空间
- 随机事件
- 随机事件的关系及其运算

E_1：掷一枚硬币观察H,T出现的情况。
E_2：将硬币抛掷三次,观察H,T出现的情况。

二. 样本空间

样本空间： 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。

例1： 写出例 $E1$～ $E6$的样本空间S。

解： $S_1=\{H,T\}$

$$S_2=\left\{\begin{array}{l}HHH,HHT,HTH,HTT,\\ THH,THT,TTH,TTT\end{array}\right\}$$

E_3：将硬币抛掷三次,观察H出现的次数。
E_4：掷一颗骰子,观察其出现的点数。
E_5：记录急救台一昼夜接到的呼唤次数。
E_6：测试灯泡的寿命

$$S_1 = \{H, T\}$$

$$S_2 = \left\{\begin{array}{l} HHH, \dots \\ THH, T\dots \end{array}\right.$$

$$S_3 = \{0, 1, 2, 3\}$$

$$S_4 = \{1, 2, 3, 4, 5, 6\}$$

$$S_5 = \{0, 1, 2, 3, \cdots 100, \cdots\}$$ 可列无穷多个样本点

$$S_6 = \{t \mid T \geq t \geq 0\}$$ 不可列无穷多个样本点

注：
- S_1---S_4有有限个样本点，S_5, S_6有无穷多个样本点。
- 样本空间中的元素是由试验目的所确定的，不同的试验目的，其样本空间是不一样的。

第一章 随机事件与概率

第一节 随机试验
第二节 样本空间及随机事件

- 随机试验
- 样本空间
- 随机事件
- 随机事件的关系及其运算

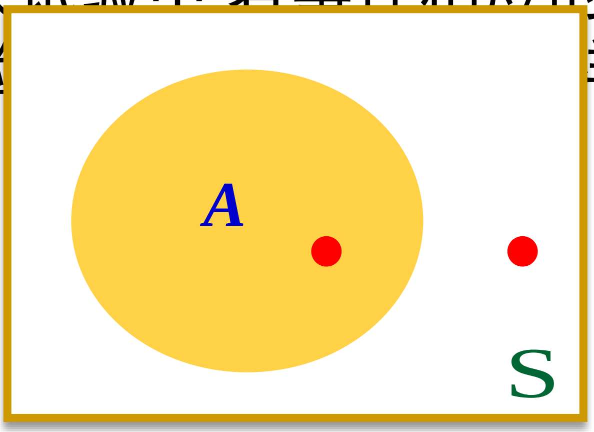
三. 随机事件

1. 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。记作A, B, C.....

2. 事件发生： 在一次试验中,当且仅当事件相应的子集中有一个样本点出现时，则称该事件发生。

注：

- 事件是人为定义的。
- 事件是否发生要进行试验：在一次试验中,若事件相应的子集中有样本点出现,则称在这次试验 … 样本点出现，则称事件没发生.



三. 随机事件

$$S_2=\left\{\begin{matrix}HHH,HHT,HTH,HTT,\\ THH,THT,TTH,TTT\end{matrix}\right\}$$

1. 随机事件： 称试验 E 的样本本空间S的子集为 E的随机事件，简称事件。

2. 事件发生： 在一次试验中,当且仅当子集中有一个样本点出现时，则称该事件发生

例2： 在 E_2 中，

事件$A_1=\{$ 第一次出现的是H $\}=\{HHH,HHT,HTH,HTT\}$

在一次试验中，事件 A_1发生有4种可能的情况.

事件$A_2=\{$ 三次出现同一面 $\}=\{HHH,TTT\}$

在一次试验中，事件A_2 发生有2种可能的情况.

三. 随机事件

3. 基本事件: 由一个样本点组成的单点集称为基本事件。

4. 必然事件： 样本空间S包含所有的样本点,它是自身的子集,在每次试验中必然发生，称为必然事件。

5. 不可能事件： 空集 $\varnothing$ 不包含任何样本点,在每次试验中都不会发生, 称为不可能事件。

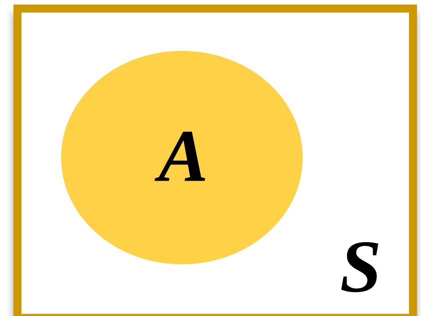
例3. 用样本点描述以下事件：

(1) 10 件产品中有一件废品，从中任取两件产品，有 一件废品。

解： 若将10件产品依次编号为1，2，….10，并设第10号产品为废品。

A = { 任取两件产品中有一件是废品 }

$$= \{(1,10),(2,10),\cdots\cdots,(9,10)\}$$

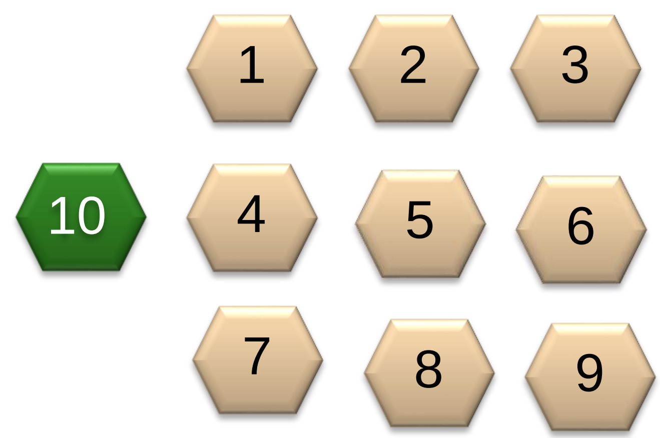


$$S = \{(1,2),(1,3),\cdots,(1,10),$$

$$(2,3),\cdots,(2,10),$$

$$\cdots\cdots\cdots\cdots$$

$$(9,10)\}$$



例3. 用样本点描述以下事件：

(2) 掷两颗骰子，点数之和小于5。

(1, 2)

(2, 1)

解： $B=\{\text{ 两颗骰子点数之和小于 5 }\}=\{(i,j)\,|\,i+j<5\}$

(i,j) 表示样本点，即两颗骰子掷出的点数

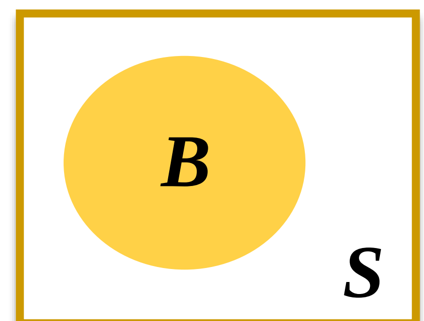
若将两颗骰子看作是不同的,分别叫做骰子A 和骰子B,

则本试验是可重复的排列问题。

$$S=\{(i,j)\,|\,i,j=1,2,\cdots,6\}$$

样本点个数=36

$$\begin{aligned}&=\{(1,1),(1,2),\cdots,(1,6),\\&\quad(2,1),(2,2),\cdots,(2,6),\\&\quad\cdots\cdots\cdots\cdots\\&\quad(6,1),(6,2),\cdots,(6,6)\}\end{aligned}$$



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

第一章 随机事件与概率

第一节 随机试验

第二节 样本空间及随机事件

- 随机试验
- 样本空间
- 随机事件
- 随机事件的关系及其运算

四. 随机事件的关系及其运算

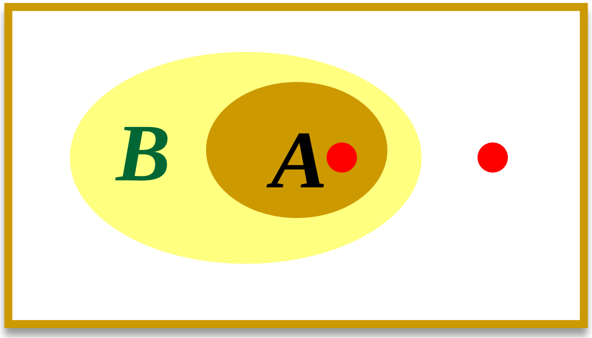
设试验 E 的样本空间为 S, 而 $A, B, A_k\ (k=1,2,\cdots)$是 S 的子集。

1. 事件的包含： 若 $A\subset B$, 则称事件B包含事件A。

→ 事件A发生必导致事件B发生

注： ➢ $A\subset B$的一个等价说法:

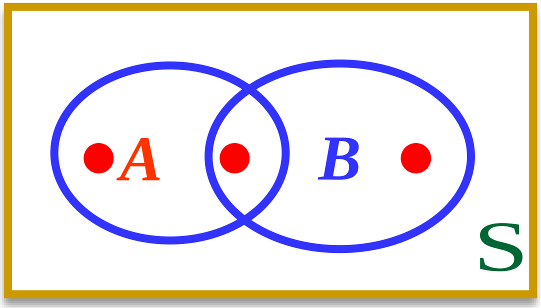
B不发生必然导致A也不发生。



$A\subset B$

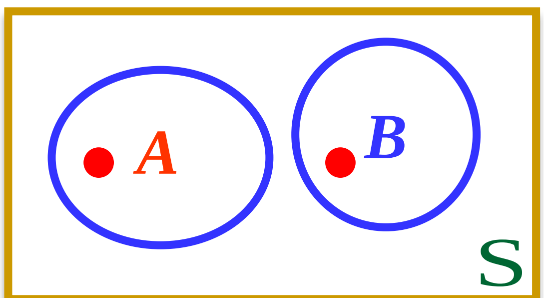
2. 事件的和(并)： 事件 $A \cup B = \{\ x \mid x \in A\ \ 或\ \ x \in B\ \}$ 称为事件 A 与事件 B 的和事件。

$A \cup B$



事件 $A \cup B$ 发生

$\Leftrightarrow$ 事件 A, B 至少有一个发生



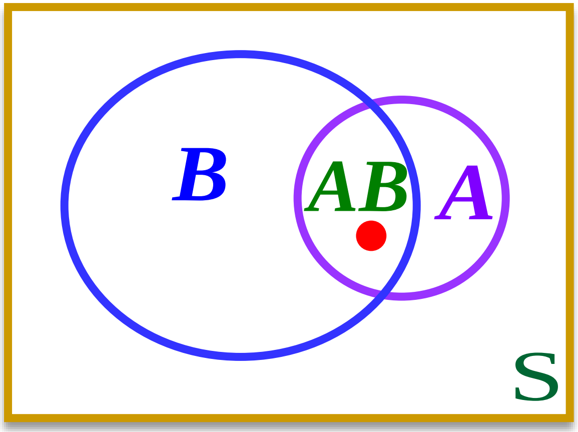
事件 $A \cup B$ 发生

$\Leftrightarrow$ 事件 A, B 有且仅有一个发生

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

3. 事件的积(交)： 事件$A\cap B=\{x \mid x\in A 且 x\in B\}$ 称为事件A与事件 B 的积(交)事件。



$A\cap B = AB$

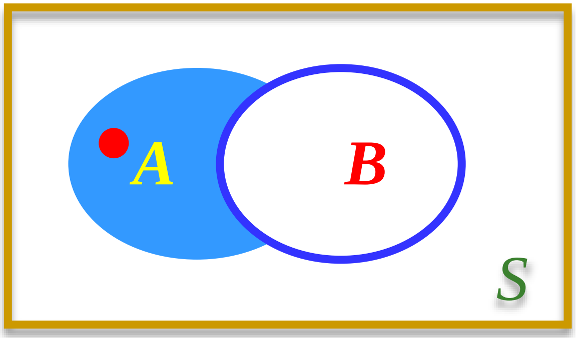
事件 $A\cap B$ 发生

$\Leftrightarrow$ 事件A与 B 同时发生

注：

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

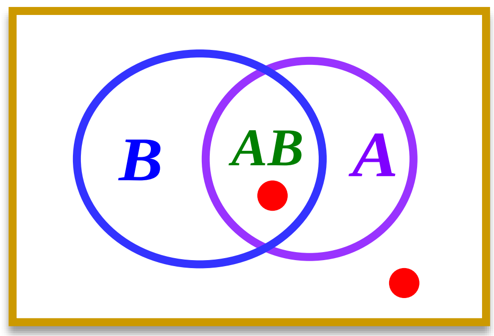
4. 事件的差： 事件 $A-B=\{x \mid x\in A且x\notin B\}$ 称为事件A与事件B的差事件。

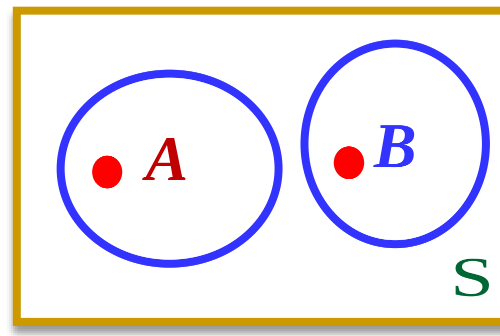


$A-B$

事件 $A-B$ 发生

$\Leftrightarrow$ 事件A 发生而事件B不发生





5. 互不相容： 若 $AB=\Phi$，则称事件A与事件B互不相容。

(互斥)事件：

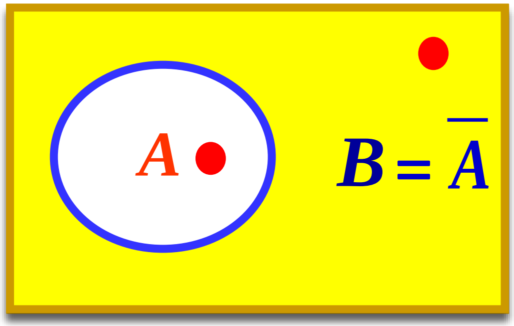
$\longrightarrow$ 事件A与事件B不同时发生

讨论：

$AB\neq\Phi$ $\not\longrightarrow$ 事件A与事件B同时发生

AB发生 $\Leftrightarrow$ 事件A与事件B 同时发生 √

6. 对立事件：（逆事件） 若 $A\cup B=S$ 且 $A\cap B=\Phi$，则称事件 A 与 B 互为对立事件。 → 事件A，B中有且仅有一个发生



A 的对立事件记为：$\overline{A}=S-A=B$

注：

▲ $A\overline{A}=\Phi,\ A\cup\overline{A}=S,\ \overline{A}=S-A,\ \overline{\overline{A}}=A$

▲ $\overline{A}$ 发生 → A 不发生

▲ 事件 $A-B$ 发生

$\Leftrightarrow$ 事件 A 发生而事件 B 不发生：$\boxed{A-B=A\overline{B}}$

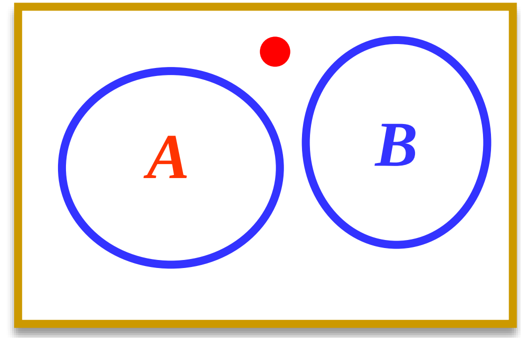
对立与互斥的联系和区别：

1）A, B对立 $\not\rightleftarrows$ 互斥

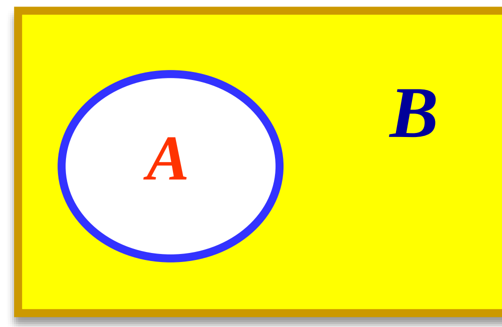
2）A, B互斥 $\longrightarrow$ 不能同时发生，但可以都不发生

A, B对立 $\longrightarrow$ 有且仅有一个发生，即肯定有一个发生

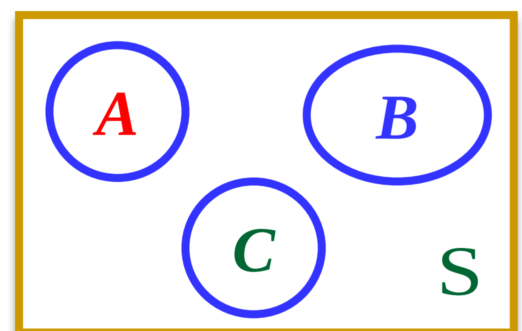
3）对立只适用于两个事件，互斥适用于多个事件。



互斥



对立



四. 随机事件的关系及其运算

复习

随机事件的三种关系：

包含： $A \subset B$ → A发生必导致B发生

互斥：$AB = \Phi$ → A与B不同时发生

对立：$A \cup B = S, A \cap B = \Phi$ → A, B中有且仅有一个发生

随机事件的三种运算：

和事件 $A \cup B$ 发生 $\Leftrightarrow$ A, B至少有一个发生 $AB \neq \Phi$

$\Leftrightarrow$ A, B有且仅有一个发生 $AB = \Phi$

积事件 AB 发生 $\Leftrightarrow$ A 与 B 同时发生

差事件 $A - B$ 发生 $\Leftrightarrow$ A 发生而 B 不发生

事件运算所满足的定律：

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

2. 结合律： $A\cup(B\cup C)=(A\cup B)\cup C$

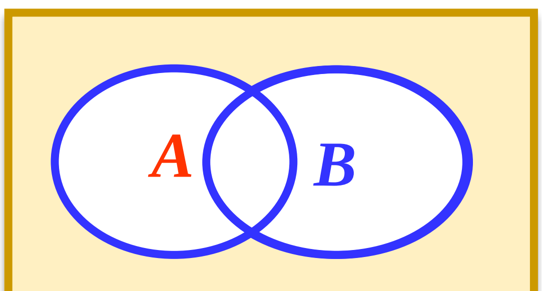
$A\cap(B\cap C)=(A\cap B)\cap C$

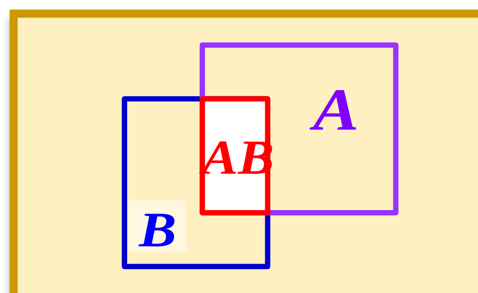
3. 分配律： $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$

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

4. **对偶定律：** $\overline{A\cup B}=\overline{A}\cap\overline{B}$ $\qquad \overline{A\cap B}=\overline{A}\cup\overline{B}$

德. 摩根律：





例4 随机试验E：对某一目标接连进行两次射击,

记 A_1 ={ 第 1 次射击命中目标 }

A_2 ={ 第 2 次射击命中目标 }

试用事件的关系和运算表示下列各事件：

(1) 第 i 次射击未命中目标, $i=1, 2$

解： 第 1 次射击未命中目标 $\overline{A_1}$

第 2 次射击未命中目标 $\overline{A_2}$

(2) $B_j=${两次射击恰好有 j 次命中目标} $j=0,1,2$

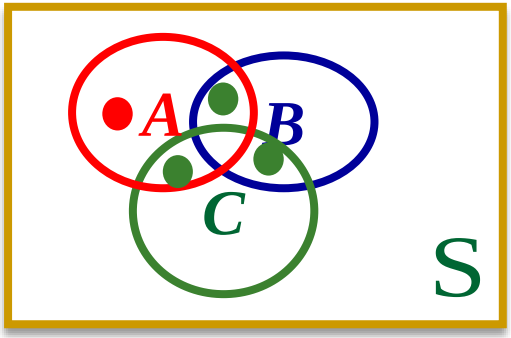
$B_0=\overline{A_1}\,\overline{A_2}$, $B_1=A_1\overline{A_2}\cup\overline{A_1}A_2$, $B_2=A_1A_2$

(3) $C_k=${两次射击至少有k次命中目标}, $k=1,2$

$C_1=A_1\cup A_2$, $C_2=A_1A_2$

例5 在某城市中共发行三种报纸：A, B, C.

试用事件的关系和运算表示下列各事件：



1）只订购报纸 A $\qquad A\bar{B}\bar{C}$

2）只订购报纸 A 与 B $\qquad AB\bar{C}$

3）只订购一种报纸 $\qquad A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$

4）正好订购两种报纸 $\qquad AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$

5）至少订购一种报纸 $\qquad A\cup B\cup C$

6）不订购任何一种报纸 $\qquad \bar{A}\bar{B}\bar{C}=\overline{A\cup B\cup C} \qquad \overline{ABC}$

解： 设事件A ={订购报纸A} $\qquad \bar{A}$={不订购报纸A}

事件B ={订购报纸B} $\qquad \bar{B}$ ={不订购报纸B}

事件C ={订购报纸C} $\qquad \bar{C}$ ={不订购报纸C}

第一章 随机事件与概率

第一节 随机试验
第二节 样本空间及随机事件

- 随机试验
- 样本空间
- 随机事件
- 随机事件的关系及其运算★

要求 熟练运用事件的关系和运算表示给出的事件

第一章 随机事件与概率

一个事件*A*在一次试验中可能发生,也可能不发生.

↓

希望知道它在一次试验中发生的**可能性大小**

↓

事件*A*发生的**概率**——$P(A)$

↓

频率 → 统计定义 → 公理化定义 → 概率的性质

第一章 随机事件与概率

第三节 频率与概率

- 频率 概率的统计定义
- 概率的公理化定义
- 概率的性质

一. 频率 概率的统计定义

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，

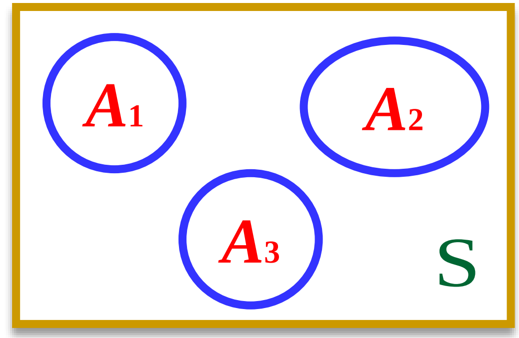
而比值 $\dfrac{n_A}{n}$ 称为事件A发生的频率。

记作：$f_n(A)=\dfrac{n_A}{n}$ —— 在n次试验中A发生的频繁程度

直观结论：$f_n(A)$越大， $P(A)$ 越大

问题： 能否用 $f_n(A)$表示 $P(A)$

2. 频率的性质：



(1) $0 \le f_n(A) \le 1$ (非负性) $f_n(A) = \dfrac{n_A}{n}$

(2) $f_n(S) = 1$ (规范性)

(3) 若 $A_1, A_2, \cdots A_k$ 是两两互斥的事件，则 (有限可加性)

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

有且仅有一个发生

3. 频率的稳定性：

在不变的条件下, 重复进行 n 次试验, 事件A发生的频率 $f_n(A)$将稳定地在某一常数 p 附近摆动, 并且 n 越大，摆动幅度越小，则称常数 p为频率的稳定值。(Ch5-大数定律)

例1：抛硬币　　设 $A=\{\text{出现正面}\}$

观察试验的结果：

序号	$n=5$		$n=50$		$n=500$	
	n_A	$f_n(A)$	n_A	$f_n(A)$	n_A	$f_n(A)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
		0.8		0.18		0.036

$f_n(A)$ 的特点：

1) 具有波动性：

$n\uparrow,\ f_n(A)$波动幅度$\downarrow$

例1： 抛硬币　设 $A=\{$出现正面$\}$

试验者	n	n_A	$f_n(A)$
德. 摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

观察试验的结果：

$f_n(A)$ 的特点：

1) 具有波动性：

$n\uparrow,\ f_n(A)$ 波动幅度 $\downarrow$

2) 具有稳定性：

$$f_n(A)\xrightarrow{n\to\infty}0.5$$

$P(A)=$ 频率的稳定值

概率的统计定义

注：统计定义的问题：

- 统计过程太麻烦；
- 统计方法只适用于简单事件。

第一章 随机事件与概率

第三节 频率与概率

- 频率 概率的统计定义
- 概率的公理化定义
- 概率的性质

柯尔莫哥洛夫，A. H.

1933年，前苏联数学家柯尔莫哥洛夫给出了概率的公理化定义.

即: 通过规定概率应具备的基本性质来定义概率.

二. 概率的公理化

设E是随机试验，

都赋予一个实数 P

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

(3) 有限可加性：$A_1, A_2, \cdots A_k$ 两两互斥，则

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

(1) 非负性： 对于每一事件A有： $0 \le P(A) \le 1$

(2) 规范性： $P(S)=1$

(3) 可列可加性：设 $A_1, A_2, \cdots A_k \cdots$ 是两两互斥的事件，

则 $P(\underbrace{A_1 \cup A_2 \cup \cdots \cup A_k \cdots}_{\text{有且仅有一个发生}}) = P(A_1) + P(A_2) + \cdots + P(A_k) + \cdots$

则称 $P(A)$ 为事件A发生的概率。

注：
- 该定义只规定了概率必须满足的最基本性质，并没有解决概率如何计算的问题。
- 该定义的意义在于它为普遍而严格的概率理论奠定了基础。

可列可加性的完备性：

(3) 可列可加性：设 $A_1, A_2, \cdots A_k \cdots$ 是两两互斥的事件，

则 $P(A_1 \cup A_2 \cup \cdots \cup A_k \cdots) = P(A_1) + P(A_2) + \cdots + P(A_k) + \cdots$

E：记录120急救台一昼夜接到的呼叫次数。

$$S = \{0, 1, 2, 3, \cdots 100, \cdots\}$$

$$= \{0\} \cup \{1\} \cup \{2\} \cup \{3\} \cup \cdots \cup \{100\} \cup \cdots$$

$$1 = P(S) = P(\{0\} \cup \{1\} \cup \{2\} \cup \{3\} \cup \cdots \cup \{100\} \cup \cdots)$$

$$= P\{0\} + P\{1\} + P\{2\} + P\{3\} + \cdots + P\{100\} + \cdots$$

可列可加性

第一章 随机事件与概率

第一节 随机试验

第二节 样本空间及随机事件

- 随机试验
- 样本空间
- 随机事件
- 随机事件的关系及其运算★

第一章 随机事件与概率

第三节 频率与概率

- 频率 概率的统计定义
- 概率的公理化定义
- 概率的性质

作 业

授课内容	习题一
1.1 随机事件1.2 样本空间	2 √
1.3 频率与概率	3(2)(3)
1.4 等可能概型	6,7,8,11等可能
1.5 条件概率	14,15, 条件概率
1.6 乘法定理，全概率，贝叶斯，独立性	17,18乘法定理 21,23,24,26全概率贝叶斯 28,29独立性

提交作业截止时间：周二晚9点