江西理工大学多媒体课件

概率论与数理统计

理学院　　数学教研室

《概率论与数理统计》

前　言

概率统计是研究随机现象数量规律的数学学科，理论严谨，应用广泛，发展迅速. 目前, 不仅高等学校各专业都开设了这门课程，而且从上世纪末开始，这门课程特意被国家教委定为本科生考研的数学课程之一，希望大家能认真学好这门不易学好又不得不学的重要课程.

主要教材及参考书

教　　材：《概率论与数理统计》第三版
江西高校出版社，1999年12月

参 考 书：《概率论与数理统计》习题精选
江西高校出版社，1999年12月

《概率论与数理统计》 (第二版)
浙江大学编　　高教出版社

《概率论与数理统计辅导》傅维潼 编　清华大学出版社　2001年

国内有关经典著作

1.《概率论基础及其应用》

王梓坤著　科学出版社　1976 年版

概率统计专业

首位中科院院士

2.《数理统计引论》

陈希儒著　科学出版社　1981年版

国外有关经典著作

1.《概率论的分析理论》

概率论的最早著作

P.- S.拉普拉斯著　1812年版

2. 《统计学数学方法》

数理统计最早著作

H. 克拉默著　1946年版

概率(或然率或几率) —— 随机事件出现的可能性的量度—— 其起源与博弈问题有关.

概率论是一门研究客观世界随机现象数量规律的 数学分支学科.

16世纪意大利学者开始研究掷骰子等赌博中的一些问题；17世纪中叶，法国数学家B. 帕斯卡、荷兰数学家C. 惠更斯 基于排列组合的方法，研究了较复杂 的赌博问题， 解决了“ 合理分配赌注问题” (即得分问题).

对客观世界中随机现象的分析产生了概率论；使 概率论 成为 数学的一个分支的真正奠基人是瑞士数学家J.伯努利；而概率论的飞速发展则在17世纪微积分学说建立以后.

第二次世界大战军事上的需要以及大工业与管理的复杂化产生了运筹学、系统论、信息论、控制论与数理统计学等学科.

数理统计学是一门研究怎样去有效地收集、整理和分析带有随机性的数据，以对所考察的问题作出推断或预测，直至为采取一定的决策和行动提供依据和建议的 数学分支学科.

统计方法的数学理论要用到很多近代数学知识，如函数论、拓扑学、矩阵代数、组合数学等等，但关系最密切的是概率论，故可以这样说：概率论是数理统计学的基础，数理统计学是概率论的一种应用. 但是它们是两个并列的数学分支学科，并无从属关系.

排列组合有关知识复习

加法原理：完成一件事情有n类方法，第i类方法中有m_i种具体的方法，则完成这件事情共有

$$\sum_{i=1}^{n} m_i$$

种不同的方法

乘法原理：完成一件事情有n个步骤，第i个步骤中有m_i种具体的方法，则完成这件事情共有

$$\prod_{i=1}^{n} m_i$$

种不同的方法

排列 从 n 个不同的元素中任意取出 m 个 (不放回地）不同的元素，按任意的顺序排成一列，称这一列为从 n 个不同的元素中取出m 个 不同的元素的一种排列。

排列数 排列的种数

$$P_n^m = n(n-1)(n-2)\cdots(n-m+1)$$

全排列 $P_n^n = n!$

可重复排列 从 n 个不同的元素中可重复地取出 m 个排成一排, 不同的排法有 n^m 种

组合 从 n 个不同的元素中任意取出 m 个 (不放回地）不同的元素，组成一组，不考虑元素的次序称为从 n 个不同的元素中取出m 个 不同的元素的一种组合。

组合数 组合的种数

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

$$\triangleq \binom{n}{m}$$

第一章　概率论的基本概念

确定性现象（决定性现象）：在一定条件下必然出现（或必然不出现）某种结果的现 象。事前可预言。

随机现象　：在一定条件下，可能出现这样的结果，也可能出现那样的结果，而且在事先无法预知确切结果的现象，事前无法预言。

1、每次试验前不能预言出现什么结果

2、每次试验后出现的结果不止一个

3、在相同的条件下进行大量观察或试 验时，出现的结果有一定的规律性

——称之为统计规律性

§1.1 随机事件、频率、概率

一、样本空间与随机事件

观察一定条件下的现象, 称为**试验.**

若它有如下特点,则称为**随机试验**,用E表示

1、 可在相同的条件下重复进行

——可重复性

2、试验结果不止一个,但能明确所有的结果

——某种意义下的确定性

3、试验前不能预知出现哪种结果

——不确定性

样本空间—— 随机试验E 所有可能的结果组成的集合称为样本空间，记为Ω

样本空间的元素, 即E 的直接结果, 称为样本点,常记为ω ，$\Omega = \{\omega\}$

随机事件 —— Ω的子集,记为 A ,B ,…

它是满足某些条件的样本点所组成的集合.

例1 给出一组随机试验及相应的样本空间

E_1：投一枚硬币3次，观察正面出现的次数

$\Omega_1 = \{0,1,2,3\}$ ⟶ 有限样本空间

E_2 :观察总机每天9:00~10:00接到的电话次数

$\Omega_2 = \{0,1,2,3,\cdots,N\}$

E_3：观察某地区每天的最高温度与最低温度

$\Omega_3 = \{(x,y) \mid T_1 \le x \le y \le T_2\}$ ⟶ 无限样本空间

其中T_1,T_2分别是该地区的最低与最高温度

基本事件 —— 仅由一个样本点组成的子集它是随机试验的直接结果,每次试验必定发生且只可能发生一个基本事件.

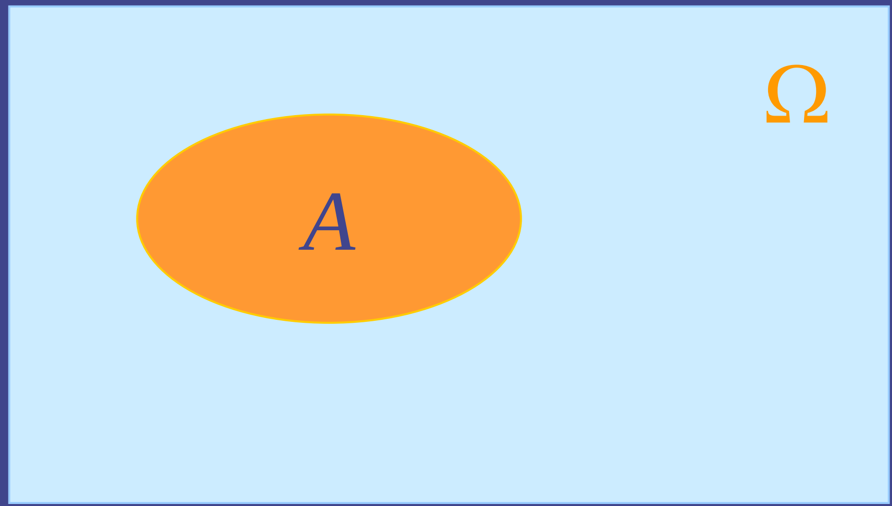
随机事件发生 —— 组成随机事件的一个样本点发生

必然事件——全体样本点组成的事件,记为Ω, 每次试验必定发生的事件.

不可能事件——不包含任何样本点的事件,记为Φ ,每次试验必定不发生的事件.

二、事件的关系和运算

Venn图

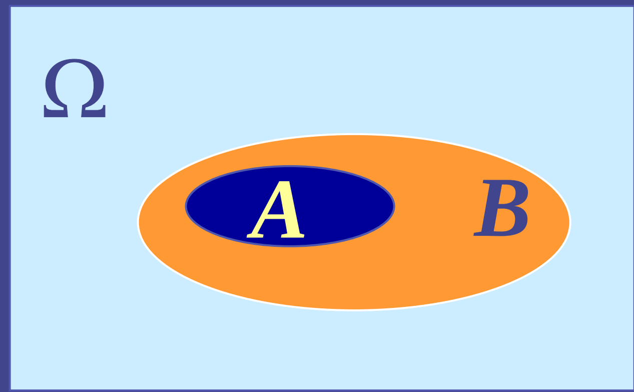


随机事件的关系和运算
雷同集合的关系和运算

1. 事件的包含

$A \subset B$ —— $\boldsymbol{A}$ 包含于$\boldsymbol{B}$

$\Leftrightarrow$ 事件 $\boldsymbol{A}$ 发生必导致事件 $\boldsymbol{B}$ 发生



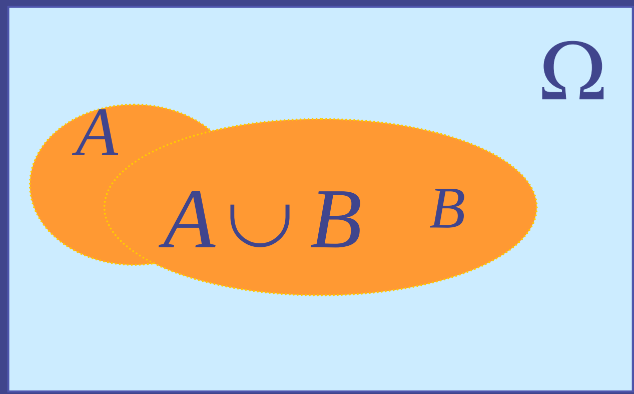
2. 事件的相等

$$A = B \Leftrightarrow A \subset B \text{ 且 } B \subset A$$

3. 事件的并(和)

$A \cup B$ 或 $A+B$

—— ***A*** 与 ***B*** 的和事件



$A \cup B$ 发生

$\Leftrightarrow$ 事件 ***A*** 与事件 ***B*** 至少有一个发生

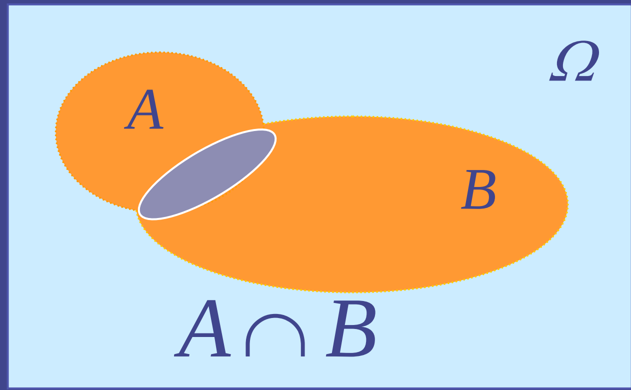
$A_1, A_2, \cdots, A_n$ 的和事件 —— $\bigcup_{i=1}^{n} A_i$

$A_1, A_2, \cdots, A_n, \cdots$ 的和事件 —— $\bigcup_{i=1}^{\infty} A_i$

4. 事件的交(积)

$A \cap B$ 或 AB

—— ***A*** 与 ***B*** 的积事件



$A \cap B$ 发生

$\Leftrightarrow$ 事件 ***A*** 与事件 ***B*** 同时发生

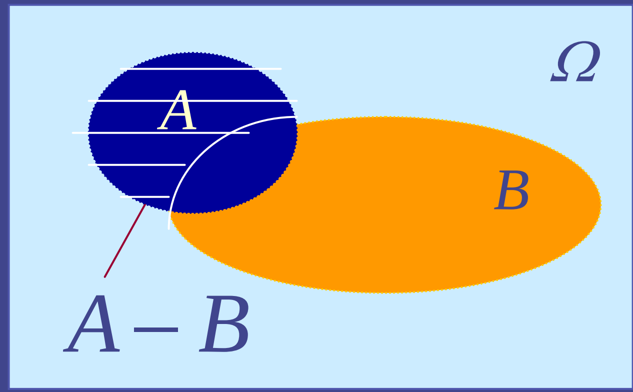
$A_1, A_2, \cdots, A_n$ 的积事件 —— $\bigcap_{i=1}^{n} A_i$

$A_1, A_2, \cdots, A_n, \cdots$ 的积事件 —— $\bigcap_{i=1}^{\infty} A_i$

5. 事件的差

$A - B$

——$\boldsymbol{A}$ 与$\boldsymbol{B}$ 的差事件

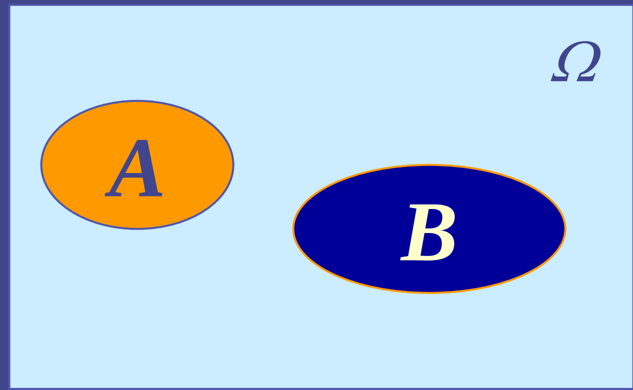


$A - B$ 发生

$\Leftrightarrow$ 事件 $\boldsymbol{A}$ 发生，但事件 $\boldsymbol{B}$ 不发生

6. 事件的互不相容（互斥）

$AB=\varnothing$ —— A 与 B 互不相容 $\Leftrightarrow$ A、B 不可能同时发生



$A_1, A_2, \cdots, A_n$ 两两互不相容

$$\Leftrightarrow A_i A_j = \varnothing, i \neq j, i, j = 1, 2, \cdots, n$$

$A_1, A_2, \cdots, A_n, \cdots$ 两两互斥(互不相容)

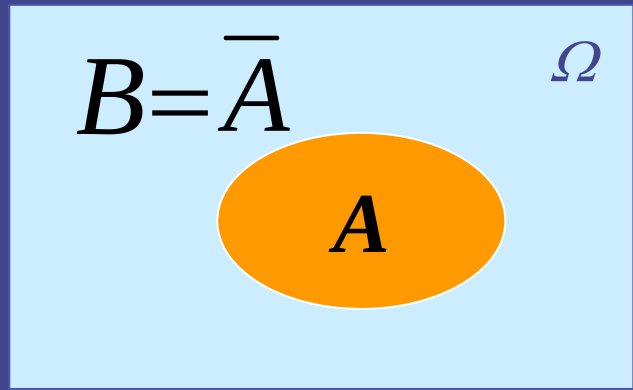
$$\Leftrightarrow A_i A_j = \varnothing, i \neq j, i, j = 1, 2, \cdots$$

7. 事件的对立

$AB=\varnothing$, $A\cup B=\Omega$

——**A 与B 互相对立**

$\Leftrightarrow$ 每次试验 **A**、 **B**中有且只有一个发生



称**B** 为**A**的对立事件(or逆事件),
记为 $B=\bar{A}$

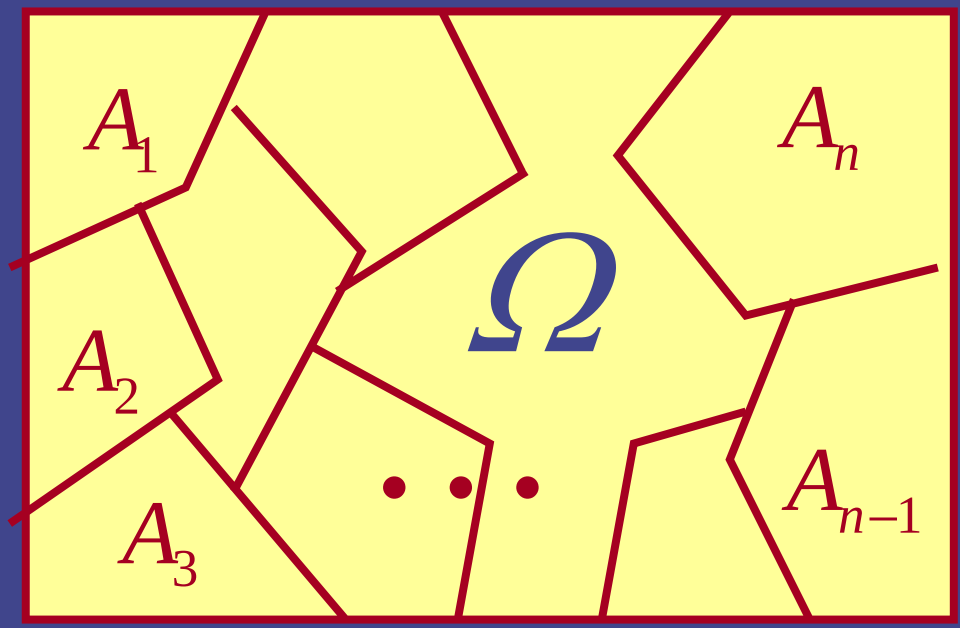
注意：“A 与 B 互相对立”与“A 与 B 互斥”是不同的概念

8. 完备事件组

若 $A_1, A_2, \cdots, A_n$ 两两互斥，且 $\Omega = \bigcup_{i=1}^{n} A_i$

则称 $A_1, A_2, \cdots, A_n$ 为完备事件组

或称 $A_1, A_2, \cdots, A_n$ 为 Ω 的一个划分



运算律

事件运算 $\xleftrightarrow{\text{对应}}$ 集合运算

- 吸收律　$A \cup \Omega = \Omega$　　$A \cap \Omega = A$

  $A \cup \varnothing = A$　　$A \cap \varnothing = \varnothing$

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

- 重余律　$\overline{\overline{A}} = A$
- 幂等律　$A \cup A = A$　　$A \cap A = A$
- 差化积　$A - B = A\overline{B} = A - (AB)$

- 交换律 $A\cup B=B\cup A$ $AB=BA$
- 结合律 $(A\cup B)\cup C=A\cup(B\cup C)$

  $(AB)C=A(BC)$
- 分配律 $(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$

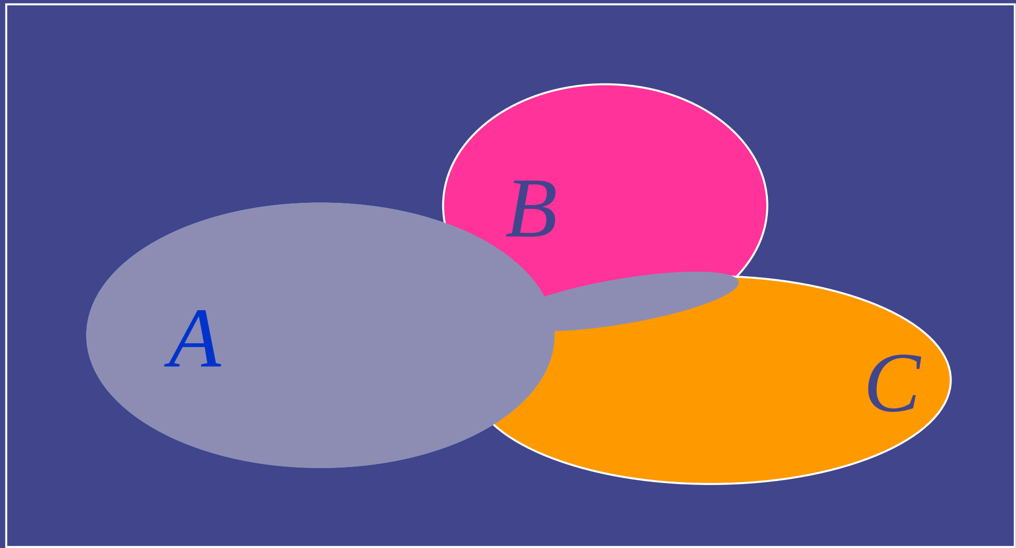
  $A\cup(BC)=(A\cup B)(A\cup C)$
- 对偶律 $\overline{A\cup B}=\overline{A}\,\overline{B}$ $\overline{AB}=\overline{A}\cup\overline{B}$

  $\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i}$ $\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}$

运算顺序：逆交并差，括号优先

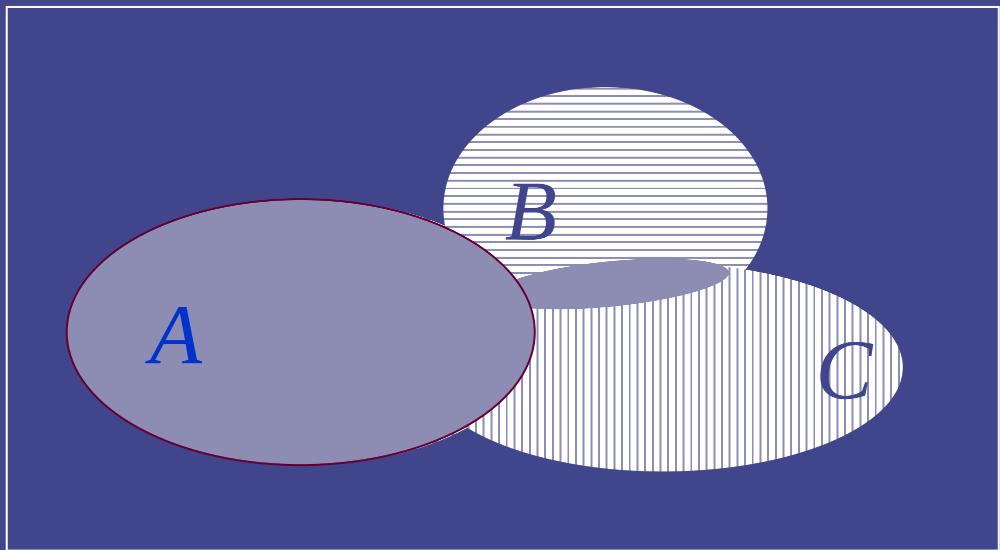
分配律

图 示

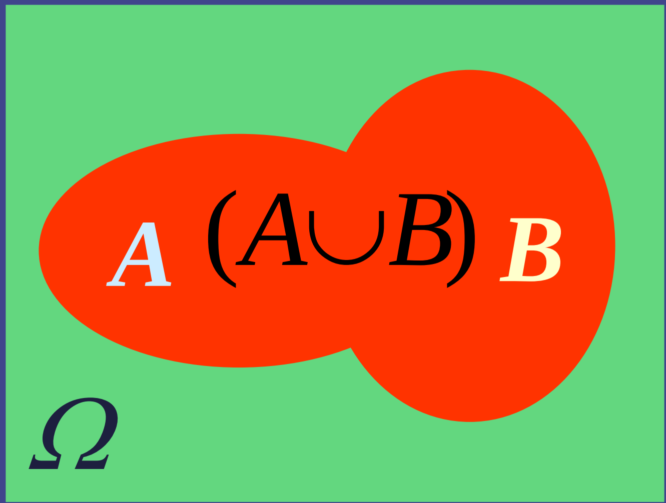


$$A \cup (BC) =$$

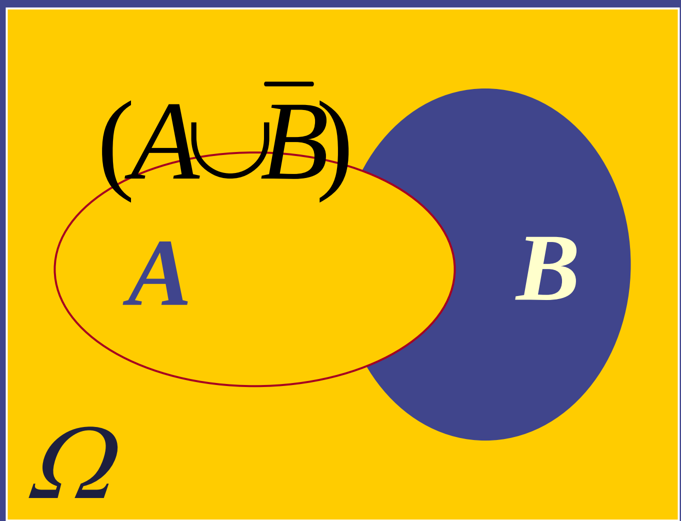
$$(A \cup B)(A \cup C)$$



例2 用图示法简化 $(A\cup B)(A\cup\bar{B})$. $AB\neq\Phi$

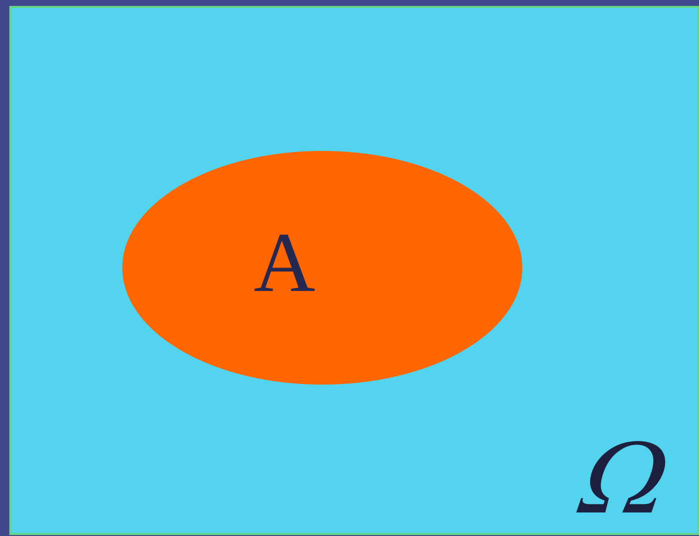


红色区域



黄色区域

交



$\therefore (A\cup B)(A\cup\bar{B})=A$

例3 化简事件 $\overline{(\overline{A}\,\overline{B}\cup C)\overline{AC}}$

解 原式 $=\overline{\overline{A}\,\overline{B}\cup C}\cup AC=\overline{\overline{A}\,\overline{B}}\,\overline{C}\cup AC$

$$=(A\cup B)\overline{C}\cup AC$$

$$=A\overline{C}\cup B\overline{C}\cup AC$$

$$=A(\overline{C}\cup C)\cup B\overline{C}$$

$$=A\Omega\cup B\overline{C}=A\cup B\overline{C}$$

例4 利用事件关系和运算表达多个事件的关系

A, B, C 都不发生——

$$\overline{A}\,\overline{B}\,\overline{C} = \overline{A \cup B \cup C}$$

A, B, C 不都发生——

$$\overline{ABC} = \overline{A} \cup \overline{B} \cup \overline{C}$$

例5 在图书馆中随意抽取一本书，

事件 A 表示数学书，

B 表示中文书，

C 表示平装书.

则

$AB\overline{C}$ —— 抽取的是精装中文版数学书

$\overline{C}\subset B$ —— 精装书都是中文书

$\overline{A}=B$ —— 非数学书都是中文版的，且中文版的书都是非数学书

三、频率和统计规律性

- 频率的定义

设在 N 次试验中，事件 A 发生了 n 次，

则称 $f_N(A)=\dfrac{n}{N}$ 为事件A在这N次试验中发生的频率

频率的性质

1、$0 \le f_N(A) \le 1$ ——— 非负性

2、$f_N(\Omega) = 1$ ——— 规范性

3、事件 $\boldsymbol{A}, \boldsymbol{B}$互斥，则

$f_N(A \cup B) = f_N(A) + f_N(B)$ ——— 可加性

可推广到有限个两两互斥事件的和事件

4、$\lim\limits_{N \to \infty} f_N(A) = P(A)$ ——— 稳定性

→ 某一定数

频率稳定性的实例

蒲丰(*Buffon*)投币——

投一枚硬币观察正面向上的次数

$N = 4040,\quad N_H = 2048,\quad f_N(H) = 0.5069$

皮尔森(*Pearson*) 投币

$N = 12000,\quad N_H = 6019,\quad f_N(H) = 0.5016$

$N = 24000,\quad N_H = 12012,\quad f_N(H) = 0.5005$

例 **Dewey G.** 统计了约438023个英语单词中各字母出现的频率, 发现各字母出现的频率不同：

A: 0.0788	B: 0.0156	C: 0.0268	D: 0.0389
E: 0.1268	F: 0.0256	G: 0.0187	H: 0.0573
I: 0.0707	J: 0.0010	K: 0.0060	L: 0.0394
M: 0.0244	N: 0.0706	O: 0.0776	P: 0.0186
Q: 0.0009	R: 0.0594	S: 0.0634	T: 0.0987
U: 0.0280	V: 0.0102	W: 0.0214	X: 0.0016
Y: 0.0202	Z: 0.0006		

概率的定义

概率的统计定义 在相同条件下重复进行的 n 次试验中, 事件 A 发生的频率稳定地在某一常数 p 附近摆动, 且随 n 越大摆动幅度越小, 则称 p 为事件 A 的概率, 记作 $P(A)$.

对本定义的评价

优点：直观易懂

缺点：粗糙模糊　不便使用

统计概率的性质

1、$0 \le P(A) \le 1$ —— 非负性

2、$P(\Omega) = 1$ —— 规范性

3、对于互不相容的事件 $A_1, A_2, \cdots, A_n$ 有

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$ —— 有限可加性