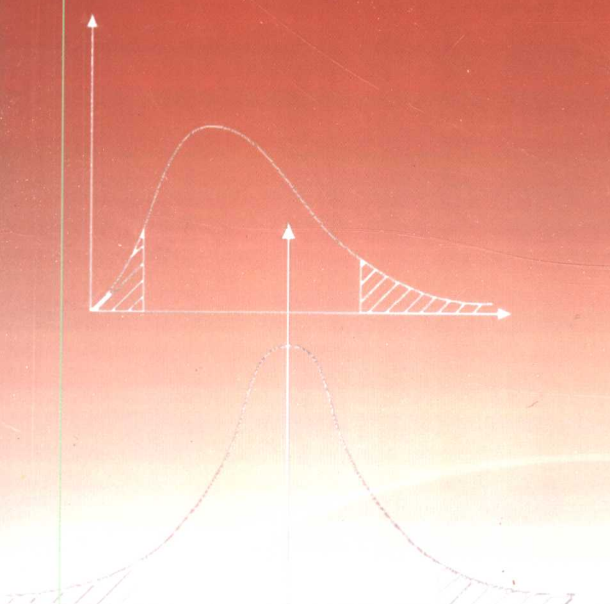
陈薇 主编

概率论与数理统计

（第2版）



CAUP

中国农业大学出版社

概率论与数理统计

（第2版）

陈　薇　主编

中国农业大学出版社

图书在版编目(CIP)数据

概率论与数理统计/陈薇主编．—2版．—北京：中国农业大学出版社，2005.8
ISBN 7-81066-886-2

Ⅰ．概…　Ⅱ．陈…　Ⅲ．①概率论 ②数理统计　Ⅳ．021

中国版本图书馆CIP数据核字(2005)第053396号

书　名　概率论与数理统计(第2版)
作　者　陈　薇　主编

策划编辑　高　欣　宋俊果　　责任编辑　杨建民　高　欣
封面设计　郑　川　　责任校对　王晓凤
出版发行　中国农业大学出版社
社　址　北京市海淀区圆明园西路2号　　邮政编码　100094
电　话　发行部 010-62731190,2620　　读者服务部 010-62732336
　　　　编辑部 010-62732617,2618　　出　版　部 010-62733440
网　址　http://www.cau.edu.cn/caup　　E-mail　caup@public.bta.net.cn
经　销　新华书店
印　刷　北京鑫丰华彩印有限公司
版　次　2005年8月第2版　2005年8月第1次印刷
规　格　787×980　16开本　16.75印张　306千字
印　数　1～3 050
定　价　21.00元

第1版前言

概率论与数理统计是高等农业院校的一门重要基础课，它不但与生物统计、田间实验设计、数量遗传、数量生态以及农业经济与管理等许多课程有密切关系，而且也是现代农业科学试验研究的重要理论基础和必需的手段、方法，因此这门课程已成为培养现代农业科学技术人才要求必须掌握的基础课程之一。

本书是以全国农业院校《概率论与数理统计》教学大纲为依据、按照农业部学科组“概率论与数理统计教学基本要求(讨论稿)”编写的。内容包括：随机事件及其概率、随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计的基本概念、参数估计、假设检验、方差分析和回归分析。在编写过程中，我们尽量做到联系农业各学科的实际问题，注重应用，力求将概念叙述得清晰易懂，做到便于教学与自学。

为了掌握和运用基本理论与公式，书中列举了较多的例题并在各章后配有一定数量的习题供课外练习。另外，为了让读者能够了解到标准化试题的类型，每章都配有少量的自测题，书后附有习题及自测题的答案供参考。本书可供50～70学时讲授，具体使用时还可根据各专业的不同情况对内容做必要的变动或取舍。

本书除可供高等农业院校作为教材外，亦可作为生物、医学等有关专业以及准备报考农业院校研究生的读者参考书。

在本书的编写过程中，得到了中国农业大学数学教研室各位老师的帮助与支持，在此表示衷心的感谢。

限于编者水平，书中缺点错漏之处诚恳地希望读者批评指正。

编　者

1997年9月

再版前言

本书在第1版的基础上，根据多年的教学经验修改而成。为适应现代数学技术的快速发展，本书还增加了Excel统计应用一章，内容包括：Excel常用统计函数、Excel在区间估计与假设检验上的应用和Excel在方差分析与回归分析上的应用。在本书修订的过程中，保持了原教材的系统和风格，充分考虑便于教学与自学的需要，叙述详细、结构严谨、结合实例、通俗易懂。值得一提的是，本书力求使读者既能获得概率论与数理统计知识，又可利用目前应用最为广泛的Office表格处理软件——Excel进行较基础的科学计算，这是在此类教材方面的初次尝试，希望能对读者的学习有所帮助。

本书由中国农业大学陈薇主编，参加编写的有：李国辉、郭雪柳、介跃建和苏时光，在此一并致谢。

编　者

2005年元月

目　　录

第一章　随机事件及其概率

§1.1　随机事件与样本空间

一、必然现象与随机现象

在自然界与人类社会中，人们会观察到各种各样的现象，把它们归纳起来，大体上可以分为两类：一类现象是可以预言其结果的，即在保持条件不变的情况下，重复进行试验其结果总是确定的．例如，水在标准大气压下加热到100℃必然沸腾；纯种紫花豌豆的后代一定开紫花；水稻从播种到收割总是经过发芽、育秧、长叶、吐穗、扬花、结实几个阶段．我们把这类现象称为必然现象或确定性现象．另一类现象是无法事先断言其结果的，即使在保持条件不变的情况下重复进行试验，其结果也未必相同．例如，在相同条件下抛同一枚硬币，其结果可能是正面朝上，也可能是反面朝上；一射手向同一目标连续射击，弹孔的位置不尽相同；在适宜条件下每穴播3粒玉米种子，有的穴可能全出苗，有的可能部分出苗，也有的可能全不出苗．这些现象在试验或观察之前都无法断言其确切的结果，具有偶然性或不确定性．那么这类现象是不是就无规律可循呢？并非如此，人们通过长期实践并深入研究之后，发现所谓不确定性，只不过是对一次或少数几次观察或试验而言的．当在相同条件下进行大量重复试验时，其试验结果却呈现出某种规律性．例如，多次重复抛一枚硬币就会发现，正面和反面出现的次数几乎各占总数的一半；一射手对同一目标进行射击时，当射击次数非常多时就可发现，弹孔是按一定规律分布的，弹孔关于目标的分布略呈对称性，且越靠近目标的弹孔越密，越远离目标的弹孔越稀．这种在个别试验中其结果呈现出偶然性，不确定性，但在大量重复试验中所得结果又呈现出固有的规律性的现象，我们称之为随机现象或偶然现象．而随机现象在大量重复试验中所呈现出的固有的规律性称为随机现象的统计规律性．概率论与数理统计就是研究和揭示随机现象统计规律性的一门数学学科．

由于随机现象广泛存在，各个不同领域所提出的大量问题促使概率统计学科蓬勃发展，使概率统计理论与方法的应用越来越广泛，几乎遍及所有科学技术领域、工农业生产和国民经济的各个部门．另外，概率统计与其他学科相结合已发展

成不少新的边缘学科，缘此我们说概率统计是许多新的重要学科的基础，更是当今农业科技工作者所必备的一种数学工具．

二、随机试验与随机事件

在概率论中首先遇到的是随机试验这个概念．为了叙述方便，我们把对在一定条件下的自然现象和社会现象所进行的观察或实验统称为试验．

一个试验如果它具有下列三个特征则称之为随机试验．

1．它是可以在相同条件下多次重复进行的试验；

2．每次试验的可能结果有多个，并且事先知道会有哪些可能的结果；

3．在进行一次试验之前，不可能事先断定哪个结果会发生．

以后我们提到的试验均指随机试验，并用E表示．下面举几个随机试验的例子．

例1.1 E_1：掷一枚硬币，观察其正反面出现的情况．

例1.2 E_2：将一枚硬币抛2次，观察其正反面出现的情况．

例1.3 E_3：在适宜的条件下播种100粒棉花种子，观察其出苗情况．

例1.4 E_4：一口袋中装有编号分别为$1,2\cdots,n$的n个球，从这袋中任取一球，观察其编号．

例1.5 E_5：掷一粒骰子，观察它出现的点数．

例1.6 E_6：测量某块地中小麦的株高．

例1.7 E_7：某一火炮向一目标射击，观察炮弹落点的分布情况．

上面举出了7个试验的例子，每一个都具有随机试验的3个特点，故它们都是随机试验．

我们把随机试验中，可能发生也可能不发生的事件，称为随机事件，简称事件，通常用大写字母A,B,C等表示．而把只包含一个试验结果的事件称为基本事件，用字母ω表示．

例如，例1.5中ω_1=“出现1点”，ω_2=“出现2点”，…，ω_6=“出现6点”等，都是试验的一种可能结果，所以它们都是基本事件．而A=“出现奇数点”，B=“出现小于4的点”，C=“出现3的倍数点”等，都是试验中可能发生也可能不发生的事情，所以它们都是随机事件．不过A,B,C与$\omega_i(i=1,2,\cdots,6)$这6个基本事件不同，它们都不是基本事件而是由试验的若干个基本事件组合而成的事件，A是由$\omega_1,\omega_3,\omega_5$组合而成的，$B$是由$\omega_1,\omega_2,\omega_3$组合而成的，$C$是由$\omega_3,\omega_6$组合而成的．分别记为$A=\{\omega_1,\omega_3,\omega_5\}$，$B=\{\omega_1,\omega_2,\omega_3\}$，$C=\{\omega_3,\omega_6\}$．事件$A,B,C$发生或不发生是通过试验的结果来确定的．例如，事件$A$是由$\omega_1,\omega_3,\omega_5$ 3个基本事件来确定的，当且仅当

$\omega_1,\omega_3,\omega_5$ 3 个基本事件之一发生时，事件 A 才发生．

我们把在试验中必然发生的事情叫做必然事件，用 Ω 表示；把不可能发生的事件称为不可能事件，用 Φ 表示．

例如，例5 中"出现的点数大于0"这一事件就是必然事件．而"出现的点数大于6"这一事件就是不可能事件．必然事件与不可能事件本来没有不确定性，即它们不是随机事件，但为了方便起见，我们把它们看作2个特殊的随机事件．

三、样本空间

为了利用点集的知识来描述随机事件，我们引进样本空间的概念．

我们将试验 E 的所有基本事件组成的集合称为 E 的样本空间．如果将样本空间也看成是事件，由于在每次试验中总有一个基本事件要发生，所以这个事件在每次试验中必定发生，因而样本空间就是必然事件．为此，我们仍用 Ω 表示样本空间．而称 Ω 中的每一个基本事件 ω 为样本点，记作 $\Omega=\{\omega\}$.

试验 E 的任一个事件，都是 E 的样本空间 Ω 中的某些样本点（基本事件）组成的集合，因而它是 Ω 的子集．

例1.8　例1.1 中 E_1 的可能结果有2个，$\omega_1=$"出现正面"，$\omega_2=$"出现反面"，则 E_1 的样本点（基本事件）有2个：ω_1 和 ω_2，那么 E_1 的样本空间 $\Omega=\{\omega_1,\omega_2\}$．若设 $A=$"出现正面"这件事，则 $A=\{\omega_1\}$是 Ω 的子集．

例1.9　例1.2 中 E_2 的可能结果有4种，$\omega_1=$（正，正），$\omega_2=$（正，反），$\omega_3=$（反，正），$\omega_4=$（反，反），则样本空间 $\Omega=(\omega_1,\omega_2,\omega_3,\omega_4)$．设事件 $A=$"一正一反"，则 $A=(\omega_2,\omega_3)$是 Ω 的子集．

例1.10　例1.4 中 E_4 的基本事件有 n 个，样本点也有 n 个，$\omega_i=$"编号为 i"($i=1,2,\cdots,n$)，则样本空间 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$，设 $A=$"编号小于10"这一事件，则 $A=\{\omega_1,\omega_2,\cdots,\omega_9\}$是 Ω 的子集．

例1.11　例1.5 中 E_5 的样本点共有6个，$\omega_i=$"出现 i 点"($i=1,2,\cdots,6$)，则样本空间 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_6\}$．而"出现3的倍数点"这一事件是2个样本点 ω_3,ω_6 组成的集合，所以它是样本空间 Ω 的子集．

例1.12　例1.6 中 E_6 若以厘米为单位，a 表示株高的下界，b 表示株高的上界，则 $\Omega=\{x\,|\,a\leqslant x\leqslant b\}$．若设事件 $A=$"株高低于 c"($a<c<b$)，则 $A=\{x\,|\,a\leqslant x<c\}$是 Ω 的一个子集．

空集不包含任何样本点，也作为样本空间的子集．若将它也看成事件的话，那么它不包含任何基本事件，故在试验中它永远不会发生．因此，空集就是不可能事件，也用 Φ 表示．

§1.2 事件的关系及运算

我们知道，在同样一组条件下往往有多个随机事件．其中有些比较简单，也有些比较复杂，并且它们往往不是孤立的，彼此之间是有联系的．因此分析事件之间的关系，认识事件的本质，想办法简化一些复杂的事件，是很有必要的．为此，我们介绍事件间的一些重要关系与事件的运算．

1．**事件的包含** 如果事件 A 发生必然导致事件 B 发生，则称事件 B 包含事件 A，或称事件 A 包含于事件 B，记作

$$B \supset A \text{ 或 } A \subset B.$$

例如，在§1.1的例1.4中，若设 A＝"编号为1或7"，B＝"编号为奇数"，则显然 $A \subset B$.

2．**事件的相等** 如果 $A \subset B$ 且 $A \supset B$，即二事件 A 与 B 中任一事件的发生必然导致另一事件的发生，则称事件 A 与事件 B 相等，记作

$$A = B.$$

例如，在§1.1的例1.4中，若设 A＝"编号不为奇数"，B＝"编号为偶数"，则显然 $A = B$.

3．**事件的和** 事件 A 与事件 B 至少有一个发生的事件称为事件 A 与事件 B 的和，记作

$$A \cup B.$$

例如，在适宜条件下，每穴播2粒玉米种子，观察出苗情况，若设 A＝"2粒都出苗"，B＝"只有1粒出苗"，C＝"至少1粒出苗"，则显然 $C = A \cup B$. 又如，§1.1例1.4中，若 A＝"编号为1"，B＝"编号为4"，则显然 $A \cup B$＝"编号为1或4".

事件的和可以推广到有限个或可数无穷多个事件的情况．

事件 $A_1, A_2, \cdots, A_n$ 中至少有一个发生的事件称为 $A_1, A_2, \cdots, A_n$ 的和，记作

$$A_1 \cup A_2 \cup \cdots \cup A_n \text{（简记为} \bigcup_{k=1}^{n} A_k\text{）}.$$

同样，事件 $A_1, A_2, \cdots, A_n, \cdots$ 中至少有一个发生的事件称为 $A_1, A_2, \cdots, A_n, \cdots$ 的和，记作

$$A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots \text{（简记为} \bigcup_{k=1}^{\infty} A_k\text{）}.$$

4. 事件的积　事件A与事件B同时发生的事件称为事件A与事件B的积(或交),记作

$$A\cap B \text{ 或 } AB.$$

例如,$A=\{\omega_1,\omega_3,\omega_5\}$,$B=\{\omega_1,\omega_2,\omega_3,\omega_4\}$,则显然$A\cap B=\{\omega_1,\omega_3\}$. 又如,§1.1例1.4中,若$A$="编号为3",$B$="编号为3的倍数",则$A\cap B=A$.

也可以将事件的积推广到有限个或可数无穷多个事件的情况.

事件$A_1,A_2,\cdots,A_n$同时发生的事件称为$A_1,A_2,\cdots,A_n$的积,记作

$$A_1\cap A_2\cap\cdots\cap A_n \text{ 或 } A_1A_2\cdots A_n(\text{简记为}\bigcap_{k=1}^{n}A_k).$$

事件$A_1,A_2,\cdots,A_n,\cdots$同时发生的事件称为$A_1,A_2,\cdots,A_n\cdots$的积,记作

$$A_1\cap A_2\cap\cdots\cap A_n\cap\cdots\text{或 } A_1A_2\cdots A_n\cdots(\text{简记为}\bigcap_{k=1}^{\infty}A_k).$$

5. 事件的差　事件A发生而事件B不发生的事件称为事件A与事件B的差. 记作

$$A-B.$$

例如,$A=\{x|10<x<100\}$,$B=\{x|50<x<200\}$,则显然$A-B=\{x|10<x\leqslant 50\}$. 由事件的差的定义可知,对于任意的事件$A$,有

$$A-A=\Phi,A-\Phi=A,A-\Omega=\Phi.$$

6. 互不相容事件　如果事件A与事件B不能同时发生,即$AB=\Phi$,则称事件A与事件B是互不相容的或互斥的.

例如,如果$A=\{x|x\geqslant 100\}$,$B=\{x|20<x<100\}$,则$AB=\Phi$,故A与B互不相容. 又如§1.1例1.4中,若设A="编号为1",B="编号为2",则A,B不能同时发生,故A与B互不相容.

7. 互逆事件　如果二事件A与B是互不相容的,并且它们中必有一事件发生,即二事件A与B中有且仅有一事件发生,即

$$A\cup B=\Omega\quad AB=\Phi$$

则称事件A与B是互逆的或对立的. 记作

$$B=\overline{A}\text{或 } A=\overline{B}.$$

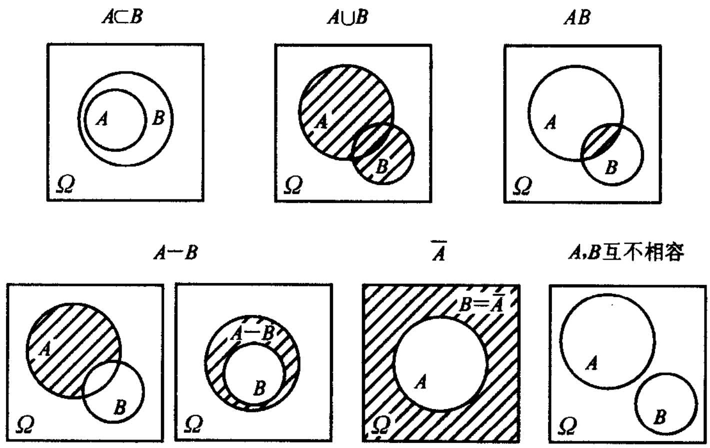
例如,设$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$,$A=\{\omega_1\}$,$B=\{\omega_2,\omega_3,\omega_4\}$,则$A\cup B=\Omega$,$AB=\Phi$,

故 A 与 B 互逆，即 $A=\overline{B}$ 或 $B=\overline{A}$．又如，§1.1 例1.4 中，若设 $A=$"编号为奇数"，$B=$"编号为偶数"，则 $A\cup B=\Omega$，$AB=\Phi$，即 $\overline{B}=A$ 或 $\overline{A}=B$，故 A，B 互为逆事件．

若 A，B 二事件互逆，则 A，B 互不相容，但反之不真．

例如，§1.1 例1.4 中，设 $A=$"编号为1"，$B=$"编号为偶数"，则 $AB=\Phi$，即 A，B 互不相容，但 $A\cup B\neq\Omega$，即 A，B 不互逆．所以 A 与 B 互不相容，不一定有 A 与 B 互逆．

事件之间的关系及运算可以用直观示意图来表示．如图 1-1 所示．



（$A\cup B$，AB，$A-B$，$\overline{A}$为图中阴影部分）

图 1-1

从上面讨论中可以看到，概率论中事件之间的关系与运算和集合论中集合之间的关系与运算完全一致．为使用方便起见，现将事件之间的关系与集合之间的关系对照如表 1-1 所示：

表 1-1

符号	概率论	集合论
Ω	样本空间，必然事件	空间（全集）
Φ	不可能事件	空集
ω	基本事件（样本点）	Ω 中的元素（或点）
A	事件	子集
$\overline{A}$	A 的对立事件	A 的余集
$A\subset B$	事件 A 发生必然导致事件 B 发生	A 是 B 的子集

续表1-1

符号	概率论	集合论
$A=B$	事件A与事件B相等	A与B相等
$A\cup B$	事件A与事件B至少有一个发生	A与B的和集
AB	事件A与事件B同时发生	A与B的交集
$A-B$	事件A发生而事件B不发生	A与B的差集
$AB=\Phi$	事件A与事件B互不相容	A与B没有公共元素

§1.3　概率的统计定义与古典概型

一、概率的统计定义

随机试验在一次试验中，其结果带有很大的偶然性，似乎没有规律可言．但在大量重复试验中，就可呈现出一定的规律性．有些事件出现的可能性大些，有些事件出现的可能性小些，有些事件出现的可能性彼此大致相同．事件出现的可能性大小，是客观存在的，它揭示了这些事件内在的统计规律．在生产实践中，了解和掌握事件发生的可能性大小具有重要意义．为了研究事件发生的可能性，需要用一个数值来描述这种可能性的大小．我们把描述这种可能性大小的数值叫做事件的概率．并用$P(A)$来表示事件A发生的概率．

对于一个给定的事件A，概率$P(A)$到底是一个什么数？怎样求？为了说明这些，先来对事件的频率下一定义．

定义1.1　若事件A在n次试验中出现μ次，则称

$$f_n(A)=\frac{\mu}{n}, \tag{1.1}$$

为事件A在n次试验中出现的**频率**．称μ为事件A在n次试验中出现的**频数**．

例1.13　蒲丰(Buffon)与皮尔逊(Pearson)曾分别掷一枚质地均匀的硬币，其结果如表1-2所示：

表1-2

试验者	掷硬币的次数	正面出现的频数	正面出现的频率
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 700	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

我们知道，掷硬币次数不多时，正面出现的次数是带有偶然性的．但表1-2说明，当试验次数充分多时，正面出现的频率在$\frac{1}{2}$左右摆动，可见出现正面次数的规律由它的频率显示出来．

例1.14 考虑某种子的发芽率．从一大批种子中抽取10批种子做发芽试验，其结果如表1-3所示：

表1-3

种子总粒数n	2	5	10	70	130	310	700	1 500	2 000	3 000
发芽种子数μ	2	4	9	60	116	282	639	1 339	1 806	2 715
种子发芽频率$\frac{\mu}{n}$	1	0.8	0.9	0.857	0.892	0.910	0.913	0.893	0.903	0.905

表1-3可以看出，发芽率在0.9附近摆动．

从上述两例的统计结果看，它们有如下特点：当我们考虑事件A发生的可能性大小时，只要我们在同一条件下做大量的重复试验，事件A发生的频率就呈现某种稳定现象．一般说来，当试验的次数增加时，事件A发生的频率总是稳定于某一数附近．频率具有稳定性这一事实，说明了描述事件A发生可能性大小的数——概率具有一定的客观存在性．

在处理实际问题时，通常我们是用试验次数足够大时的频率来度量概率的．

定义1.2 在不变的一组条件S下，重复进行n次同一试验时，如果随机事件A发生的次数为μ，当试验的次数n很大时，频率$\frac{\mu}{n}$稳定地在某一固定数值p的附近摆动；而且一般说来这种摆动幅度随着试验次数的增多会愈来愈小，则称数值p为随机事件A的**概率**，记作

$$P(A)=p.$$

在重复试验中，事件A发生的频率在$P(A)$附近摆动．因此，频率可以作为$P(A)$的一个估计值．

由频率出发所定义的事件A的概率常称为**统计概率**．所以定义1.2也就称为概率的统计定义．

当试验次数n固定时，事件A发生的频率$f_n(A)$有如下性质：

(i) $0\leqslant f_n(A)\leqslant 1$；

(ii) $f_n(\Omega)=1$；

(iii)设$A_1,A_2,\cdots,A_k$是两两互不相容的事件，则

$$f_n(\bigcup_{i=1}^{k}A_i)=\sum_{i=1}^{k}f_n(A_i).$$

证明

(i)因为 $0\leqslant\mu\leqslant n$,所以有 $0\leqslant\dfrac{\mu}{n}\leqslant1$,即有 $0\leqslant f_n(A)\leqslant1$;

(ii)因为 Ω 为必然事件,所以 $\mu=n$,故 $f_n(\Omega)=1$;

(iii)设 $A_1,A_2,\cdots,A_k$ 在 n 次试验中分别出现 $\mu_1,\mu_2,\cdots,\mu_k$ 次,由于 $A_1,A_2,\cdots,A_k$ 两两互不相容,它们不能两两同时出现,因而事件 $\bigcup_{i=1}^{k}A_i$ 在 n 次试验中出现的次数为 $\sum_{i=1}^{k}\mu_i$,故

$$f_n(\bigcup_{i=1}^{k}A_i)=(\sum_{i=1}^{k}\mu_i)/n=\sum_{i=1}^{k}\frac{\mu_i}{n}=\sum_{i=1}^{k}f_n(A_i).$$ 证毕.

概率的统计定义是在总结统计资料的基础上给出的,反映了概率的统计性质,比较直观. 但是,用这个定义的想法计算概率不方便. 因为,在实际问题中,我们不可能对每一个事件都做大量的试验,从中得到频率的稳定值. 另外,试验的次数该取多少为足够大,而且当试验的次数 n 很大时,很难保证试验的条件都完全一样. 这是统计概率的不足之处.

对于某些特殊类型问题的概率可用下面介绍的古典概率来定义.

二、古 典 概 型

定义 1.3 如果随机试验满足以下两个条件:

(i)样本空间中只有有限多个样本点,即

$$\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\};$$

(ii)每个基本事件(样本点)发生的可能性是相等的,即

$$P(\omega_i)=\frac{1}{n}\qquad i=1,2,\cdots,n$$

就称其为等可能概型试验. 等可能概型试验是概率论发展初期研究的主要对象,因此,也称为古典概型试验,简称为**古典概型**.

下面给出古典概型概率的定义.

定义1.4 设随机试验 E 为古典概型,若样本空间中所含的基本事件总数为 n,而事件 A 中所含的基本事件个数为 m,则称 $\dfrac{m}{n}$ 为随机事件 A 的概率,记作