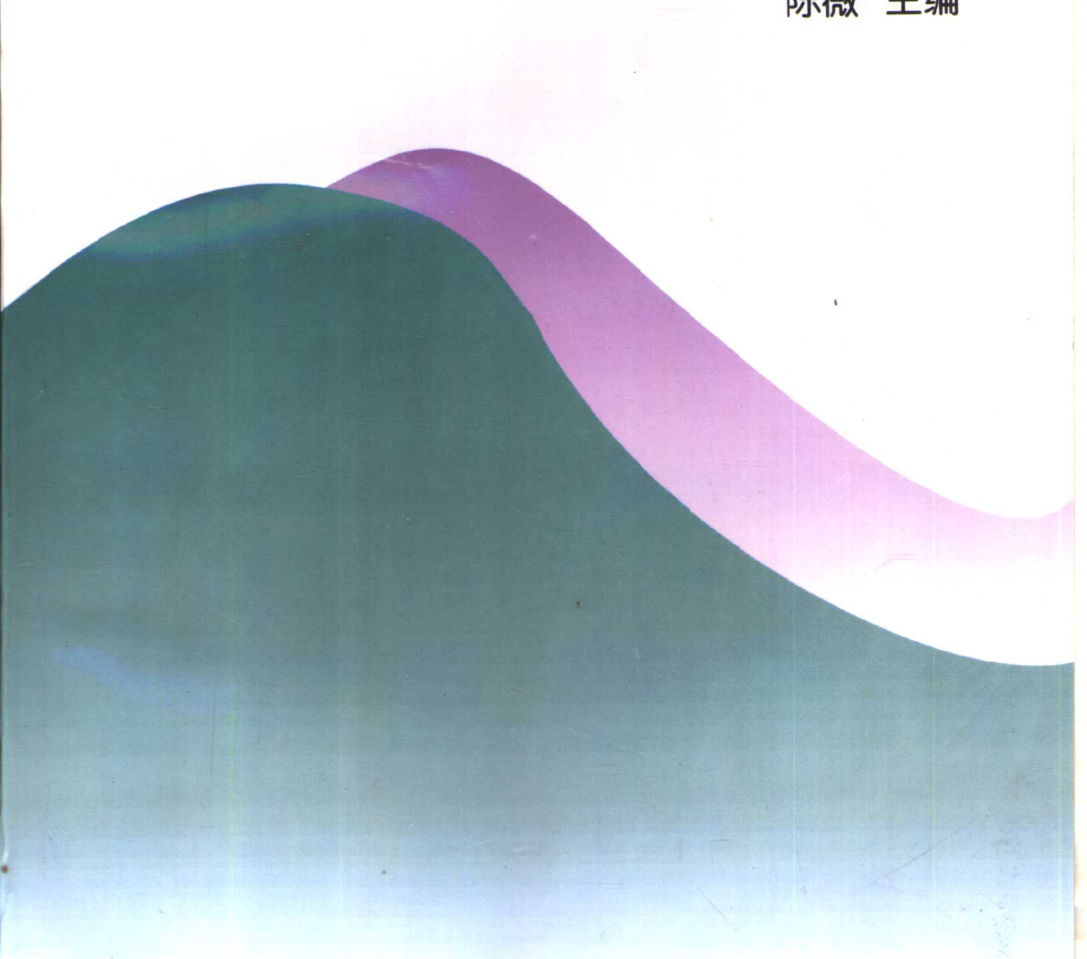


GAI Lǚ LUN YU SHU LI TONG JI

概率论与数理统计

陈薇 主编



中国农业大学出版社

概率论与数理统计

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. —北京:中国农业大学出版社,1998.3

ISBN 7-81002-873-1

I. 概… I. 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(97)第 28897 号

出 版 行 中国农业大学出版社
经 销 新华书店
印 刷 北京丰华印刷厂印刷
版 次 1998 年 2 月第 1 版
印 次 1999 年 6 月第 2 次印刷
开 本 32 印张 8.75 千字 216
规 格 850×1168
印 数 5001—10000
定 价 12.00 元

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支持与帮助,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

113000

主 编 陈 薇(中国农业大学)
副主编 李国辉(中国农业大学)
参 编 介跃建(中国农业大学)
严 瑶(中国农业大学)
郭雪柳(中国农业大学)
主 审 裴鑫德(中国农业大学)

责任编辑 宋俊果
封面设计 郑 川

目 录

第一章 随机事件及其概率	(1)
§ 1.1 随机事件与样本空间	(1)
§ 1.2 事件的关系及运算	(5)
§ 1.3 概率的统计定义与古典概型	(9)
§ 1.4 概率的一般定义及其基本性质	(18)
§ 1.5 条件概率 乘法定理	(22)
§ 1.6 全概率公式与贝叶斯公式	(26)
§ 1.7 随机事件的独立性	(30)
习题一	(34)
自测题一	(39)
第二章 随机变量及其分布	(41)
§ 2.1 随机变量及其分布函数	(41)
§ 2.2 离散型随机变量	(45)
§ 2.3 连续型随机变量	(53)
§ 2.4 随机变量函数的分布	(65)
§ 2.5 二维离散型随机变量	(69)
§ 2.6 二维连续型随机变量	(77)
习题二	(85)
自测题二	(88)
第三章 随机变量的数字特征	(91)
§ 3.1 数学期望	(91)
§ 3.2 方差	(96)
§ 3.3 二维随机变量的数字特征	(101)
§ 3.4 协方差与相关系数	(104)
§ 3.5 矩与相关矩阵	(108)

习题三	(111)
自测题三	(114)
第四章 大数定律与中心极限定理	(116)
§ 4.1 大数定律	(116)
§ 4.2 中心极限定理	(122)
习题四	(128)
自测题四	(129)
第五章 数理统计的基本概念	(131)
§ 5.1 总体与样本	(131)
§ 5.2 理论分布与经验分布	(133)
§ 5.3 统计量及其分布	(135)
习题五	(150)
自测题五	(151)
第六章 参数估计	(153)
§ 6.1 点估计	(153)
§ 6.2 区间估计	(163)
习题六	(172)
自测题六	(173)
第七章 假设检验	(175)
§ 7.1 假设检验的基本概念	(175)
§ 7.2 单个正态总体的假设检验	(179)
§ 7.3 两个正态总体的假设检验	(184)
习题七	(193)
自测题七	(194)
第八章 方差分析	(196)
§ 8.1 单因素试验的方差分析	(196)
§ 8.2 双因素试验的方差分析	(204)
习题八	(215)
自测题八	(217)
第九章 回归分析	(219)
§ 9.1 一元线性回归分析	(219)

§ 9.2 多元线性回归分析	(231)
习题九	(236)
自测题九	(237)
附表 1 几种常用的概率分布	(239)
附表 2 标准正态分布表	(240)
附表 3 泊松分布表	(241)
附表 4 t 分布表	(243)
附表 5 χ^2 分布表	(244)
附表 6 F 分布表	(246)
习题答案	(255)
主要参考书	(267)

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°C 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寻呢?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,只不过是就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，概率统计与其它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 : 将一枚硬币抛两次，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 $1, 2, \dots, n$ 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三个特点, 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 $\omega_1 = \text{“出现 1 点”}$, $\omega_2 = \text{“出现 2 点”}$, $\dots$, $\omega_6 = \text{“出现 6 点”}$ 等, 都是试验的一种可能结果, 所以它们都是基本事件. 而 $A = \text{“出现奇数点”}$, $B = \text{“出现小于 4 的点”}$, $C = \text{“出现 3 的倍数点”}$ 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 $\omega_i (i=1, 2, \dots, 6)$ 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 ω_3, ω_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 $\omega_1, \omega_3, \omega_5$ 三个基本事件来确定的, 当且仅当 $\omega_1, \omega_3, \omega_5$ 三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,我们把它们看作两个特殊的随机事件.

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 的可能结果有两个, $\omega_1 =$ “出现正面”, $\omega_2 =$ “出现反面”,则 E_1 的样本点(基本事件)有两个 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 的可能结果有四种, $\omega_1 =$ (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 ”($i = 1, 2, \dots, n$), 则样本空间 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, 设 $A =$ “编号小于 10”这一事件, 则 $A = \{\omega_1, \omega_2, \dots, \omega_9\}$ 是 Ω 的子集.

例 1.11 例 1.5 中 E_5 的样本点共有六个, $\omega_i =$ “出现 i 点”($i = 1, 2, \dots, 6$), 则样本空间 $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$. 而“出现 3 的倍数点”这一事件是两个样本点 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 $\Omega = \{x | a \leq x \leq b\}$. 若设事件 $A = \text{“株高低于 } c\text{”}$ ($a < c < b$), 则 $A = \{x | a \leq x < c\}$ 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 $\emptyset$ 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 $A = \text{“编号为 1 或 7”}$, $B = \text{“编号为奇数”}$,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 $A = \text{“编号不为奇数”}$, $B = \text{“编号为偶数”}$,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,每穴播两粒玉米种子,观察出苗情况,若

设 A = “两粒都出苗”, B = “只有一粒出苗”, C = “至少一粒出苗”, 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”.

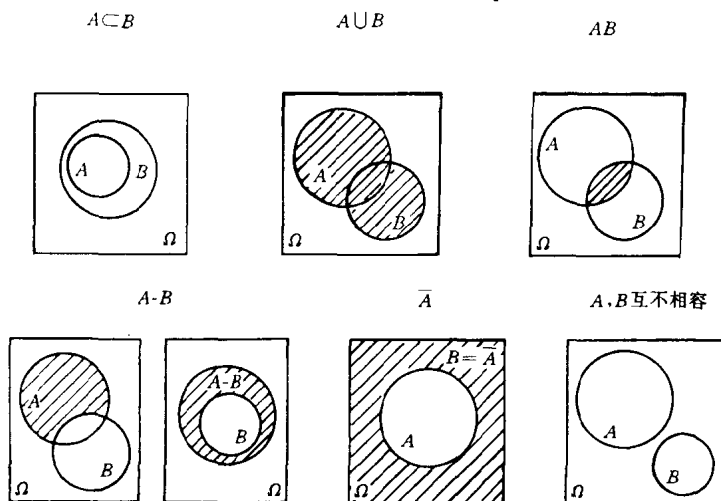
事件的和可以推广到有限个或可数无穷多个事件的情况.

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

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

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

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



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

图 1.1

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

$A \cap B$ 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 $A_1, A_2, \dots, A_n$ 同时发生的事件称为 $A_1, A_2, \dots, A_n$ 的积, 记作

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

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

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

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, $A = \{x \mid 10 < x < 100\}$, $B = \{x \mid 50 < x < 200\}$, 则显然 $A - B = \{x \mid 10 < x \leq 50\}$. 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 $A = \{x \mid x \geq 100\}$, $B = \{x \mid 20 < x < 100\}$, 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

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

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 $A = \bar{B}$ 或 $B = \bar{A}$. 又如, § 1.1 例 1.4 中, 若设 $A = \text{“编号为奇数”}$, $B = \text{“编号为偶数”}$, 则 $A \cup B = \Omega$, $AB = \Phi$, 即 $\bar{B} = A$ 或 $\bar{A} = B$, 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 $A = \text{“编号为 1”}$, $B = \text{“编号为偶数”}$, 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.

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:

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
$\bar{A}$	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
$A = B$	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

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

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:

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