概率论与数理统计

主　编　印凡成　夏乐天
副主编　袁永生　董祖引　朱永忠

河海大学出版社

内容提要

本书共分九章。包含随机事件与概率、离散型随机变量及其分布、连续型随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、样本及抽样分布、参数估计、假设检验、回归分析与方差分析。读者只需具备高等数学和线性代数的知识，就可阅读通篇内容。读者对象为大学理工科二年级以上学生和从事相关工作的科技工作者。本书在连续型随机变量的离散化处理上颇具匠心，还介绍了随机水文学中的一些统计方法，尤其是每章习题的配置，很符合教学的实际需要，具有一定的特色。

前　言

《概率论与数理统计》（简称“概率统计”）课程是理工科专业一门重要的基础理论课，其传统的教学内容是先介绍古典概率，然后分别介绍离散型随机变量和连续型随机变量及其概率分布，最后介绍以统计推断为主的数理统计。一般的教材也都是按这种思路编写的。纵观全局，有种感觉：离散型随机变量与连续型随机变量这两部分似乎是两大块关系不甚密切的内容，教学上有重复处理之嫌，但又不可不分别处理。既耗费学时，又与当前的教学改革思想不符。针对传统教材的不足之处，从1995年上半年起，应用数学教研室的一批长期从事该课程教学的教师，根据我校的具体情况，在大量查阅国内外文献的基础上，编写了这本改革试点教材。1996年9月起在校内印刷发行，主要用于综合改革试点班。使用三年多来，反映良好，现决定正式出版。

本书共分九章，有如下两个特点：

(1) 在讲清讲透离散型随机变量及其分布的前提下，通过贝努利概型和泊松呼唤流模型很自然地推出了一系列常见的分布，并把离散型、连续型的常见分布有机地联系在一起，突出了离散化处理的思想；在引入 *R-S* 积分的概念之后，给出了一般随机变量数学期望的定义，且能包含离散型连续型随机变量两种情况作为特例。

(2) 理顺了现行教材中不合理的部分，特别是理顺了习题与相应章节的衔接，力求进一步地体现理论与实践的结合。

本书第一章由袁永生编写，第二章～第四章由印凡成编写，第五章由朱永忠编写，第六章～第八章由夏乐天编写，第九章由董祖引编写，全书由印凡成、夏乐天统稿。何才法教授仔细地审阅了本书的初稿，并提出了很多建设性的意见，刘景麟教授在本书的编写过程中，自始至终给予了亲切的关怀和热情的支持，并在随机变量的离散化处理上给予了我们启发性的指导，我们在此向两位教授表示衷心的感谢。另外，在本书的出版过程中，也得到了河海大学教务处和河海大学出版社领导的大力支持和帮助，在此表示真诚的感谢。最后，我们希望广大读者和有关专家对本书提出宝贵的意见，指出其错误和缺点，以便进一步改进。

在本书的修订过程中，黄健元、徐小明老师也认真地阅读了全书，并提出许多有价值的意见，我们特在此一并表示真诚的感谢！

编　者

1999年11月于河海大学

2001年11月修订再版

目　　录

第一章　随机事件与概率

概率论有着悠久的历史,它的起源与赌博问题有关。16 世纪,意大利学者开始研究掷骰子等现象中的一些简单问题,例如比较掷两个骰子出现点数之和为 9 与 10 的可能性大小。17 世纪中叶,法国数学家帕斯卡(B.Pascal)、费尔马(P.de Fermat)及荷兰数学家惠更斯(C.Huygens)基于排列组合方法,研究了一些较复杂的赌博问题,他们解决了"分赌注问题"、"赌徒输光问题"等。

随着 18、19 世纪科学的发展,人们注意到在某些生物、物理和社会现象与机会游戏之间有一种相似,从而由机会游戏起源的概率论被应用到这些领域中,同时也大大推动了概率论本身的发展。

我们今天所知道的概率的数学理论,起源于较近代。1933 年前苏联数学家柯尔莫哥洛夫在他的主要著作《概率论基础》一书中把概率公理化了,这不仅给概率论提供了一个逻辑上的坚实基础,同时也把它与现代数学的主要倾向联系了起来,使概率论成为严谨的数学分支,对近几十年概率论的迅速发展起了重大的作用。

数理统计是概率论的一个姐妹学科,研究怎样有效地收集、整理和分析带有随机性质的数据,以对所观察的问题作出推断和预测,直至为采取一定的决策和行动提供依据和建议。

数理统计学是随着概率论的发展而发展起来的。当人们认识到必须把数据看成来自具有一定概率分布的总体,所研究的对象是这个总体而不能局限于数据本身时,数理统计也就随之诞生了。数理统计发展成一门成熟的学科,是在 20 世纪上半叶,以 1946 年克拉默(H.Cramér)发表的《统计学的数学方法》为标志。

数理统计学用到很多近代数学知识,但与其关系最密切的是概率论。可以说,概率论是数理统计的理论基础,数理统计是概率论的一种应用,并且补充和丰富了概率论。

目前,概率论与数理统计的理论与方法已广泛地应用于自然科学、技术科学、社会科学及人文科学的各个领域。例如,使用概率统计方法可以进行气象预报、地质预报及产品的抽样验收,可以给出元件或系统的使用可靠性及平均寿命的估计等等。

随着计算机的发展与普及,概率论与数理统计已成为处理信息、制定决策的重要理论和方法。概率论与数理统计向各个领域渗透,产生了许多新的分支和边缘科学,如生物统计、统计物理、教育统计等等,同时概率论与数理统计又是许多新的重要学科的基础,如信息论、控制论、排队论、预测论、可靠性理论及人工智能等。

概率论与数理统计,作为理论严谨、应用广泛、发展迅速的数学分支正日益受到人们的重视并发挥着重大的作用。

§1.1　随机现象与统计规律

概率论与数理统计是研究随机现象数量规律的一门数学分支。什么是随机现象呢?让我们先看两个现象。

Ⅰ:一个盒子中有 10 个完全相同的白球,搅匀后从中任意摸取一球;

Ⅱ：一个盒子中有 10 个形状相同的球，5 个是白色的，另外 5 个是黑色的，搅匀后从中任意摸取一球。

对于Ⅰ，在球没有取出之前，我们就能确定取出的必定是白球。这种现象根据开始时的条件就可以确定最后的结果。类似的现象很多，如在标准大气压下，纯水加热到 100℃时必然会沸腾等等，这类现象称为确定性现象；对于Ⅱ，在球没有取出之前，从开始时的条件，不能确定最后的结果是白的还是黑的，骤然一看似乎没什么规律可言。类似的现象也很多，如记录一个寻呼台每天上午 8:00～9:00 收到的用户寻呼次数，可能是 0 次、1 次、2 次、…；某地区的年降雨量等也是这类现象。这些现象的特点是：在基本条件不变的情况下，一系列试验或观察会得到不同的结果，每次试验或观察之前，不能完全确定会出现哪种结果。但是，人们经过长期实践并深入研究之后，发现这类现象在大量重复试验或观察下，它的结果却呈现某种规律性。例如多次重复抛一枚硬币得到的正面朝下的次数大致有一半；同一门炮射击同一目标的弹着点按照一定规律分布等等，这种现象称之为随机现象、其中所隐含的规律性称为统计规律。

概率论与数理统计就是研究随机现象的统计规律的数学学科。概率论研究随机现象有其独特方法，它不是企图去寻找出现每一结果的一切物理因素，从而像确定性现象那样确定无疑地预报在哪些条件下出现某一确定的结果，而是通过对随机现象的大量观察，揭示其规律性。由于随机现象的普遍性，使得概率论与数理统计具有极其广泛的应用。

§1.2 样本空间、随机事件

一、样本空间

我们先讨论一些简单的问题。

例 1.1 掷一枚硬币（假设硬币的构造是均匀的），观察正面 H、反面 T 出现的情况。

由于硬币构造均匀，因而出现“正面向上”与“反面向上”的机会均等，又无其它可能，所以我们断言，掷一枚硬币，掷得“正面向上”的可能性为 $\frac{1}{2}$。

例 1.2 将一枚硬币连抛三次，观察正面出现的次数。

例 1.3 将一枚硬币连抛三次，观察正面 H、反面 T 出现的情况。

例 1.4 掷一颗骰子，观察出现的点数。

以上四例中所进行的试验有着一些共同的特性：

1° 可以在相同的条件下重复进行；

2° 所有可能结果事先明确，并且不止一个；

3° 一次试验之前，不能肯定这次试验会出现哪一个结果。

概率论中称具有以上三特点的试验为随机试验，也简称为试验，常用 E 表示；称随机试验的一切可能结果所组成的集合为此试验的样本空间，常用 Ω 表示；称随机试验的每一个可能结果，即样本空间的元素为样本点，常用 ω 表示。

例 1、2、3、4 中试验的样本空间分别为：

$\Omega_1=\{H,T\}$；

$\Omega_2=\{0,1,2,3\}$；

$\Omega_3=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}$；

$\Omega_4 = \{1,2,3,4,5,6\}$。

这里有个问题需要注意：例 1.2 与例 1.3 同为"一枚硬币连抛三次"，由于试验的目的不同，它们样本空间的元素也不同，即样本空间的元素由试验的目的确定。

二、随机事件

在随机试验中，有时人们只关心带有某些特征的现象是否发生。如在例 1.4 中我们可以研究 $A=\{$点数$>5\}$，$B=\{$出现偶数点$\}$，$C=\{$点数$\leqslant 2\}$这样一些结果是否发生。它们都是由例 1.4 样本空间的多个样本点组成，是相应样本空间的子集。显然，这些结果的发生与否在试验之前也具有不确定性。

一般将试验 E 的样本空间 Ω 的子集称为 E 的随机事件，简称事件，习惯上常用大写字母 A、B、C 等表示。在每次试验中，如果出现 A 中所包含的某个样本点 ω，则称作 A 发生。特别地，称由一个样本点组成的单点集为基本事件，有时亦用 ω 简记。

我们已经知道，样本空间 Ω 包含了全体样本点，因而在任一次试验中，必然要出现 Ω 中的某一样本点。即在每次试验中，Ω 必然会发生，称为必然事件，仍用 Ω 表示。相应地，空集$\varnothing$也是 Ω 的子集，在每次试验中不可能出现$\varnothing$中的样本点。事实上，它不包含任何样本点，称为不可能事件，用$\varnothing$表示。必然事件和不可能事件的发生与否已经失去了"不确定性"，因此在本质上，它们不是随机事件。当然为方便起见，我们还是把它们看作随机事件，只不过是随机事件的两个极端情形而已。

一个样本空间 Ω 中，可以有很多的随机事件，概率论的任务之一就是研究随机事件本身的客观规律，进一步地通过对较简单事件规律的研究去掌握更复杂事件的规律。因此，需要研究事件之间的关系和事件之间的一些运算。

三、事件之间的关系及其运算

设试验 E 的样本空间为 Ω，而 $A,B,A_i(i=1,2,\cdots,n)$是 Ω 中的一些事件。

1° 如果事件 A 发生必然有事件 B 发生，则称事件 B 包含事件 A，记为 $A\subset B$。特别地，若 $A\subset B$，且 $B\subset A$，则称事件 A 与事件 B 相等，记为 $A=B$。如前面提到的例 1.4 中 $A=\{$点数$>5\}=\{6\}$发生就有 $B=\{$出现偶数点$\}=\{2,4,6\}$的发生。

包含的含意可以给一个直观的几何解释：设样本空间 Ω 是一个正方形(如图 1-1)，A 与 B 是两个事件，即 Ω 的两个子集。"A 发生必然有 B 发生"意味着"属于 A 的 ω 必然属于 B"，即 A 中的元素(样本点)全在 B 中。由此可见，事件 $A\subset B$ 的含意与集合论中的意义是一致的。

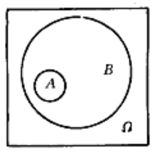

图 1-1

2° "事件 A 与事件 B 同时发生"这一事件称为事件 A 与事件 B 的积事件。记为 $A\cap B=\{\omega\mid\omega\in A$，且 $\omega\in B\}$，$A\cap B$ 也记为 AB。这也可以给它一个几何解释，如图 1-2 中阴影部分。

由于组成 $A\cap B$ 的样本点既在 A 中又在 B 中。反之，既出现在 A 中又出现在 B 中的样本点，必出现在 $A\cap B$ 中，所以事件 $A\cap B$ 发生$\Leftrightarrow A,B$ 同时发生。

同样，称$\bigcap\limits_{i=1}^{n}A_i$ 为 n 个事件 $A_1,A_2,\cdots,A_n$ 的积事件；称$\bigcap\limits_{i=1}^{\infty}A_i$ 为可列个事件 $A_1,A_2,\cdots$的积事件。

3° "事件 A 与事件 B 中至少有一个发生"这一事件称为事件 A 与事件 B 的和事件，记为 $A\cup B=\{\omega\mid\omega\in A$ 或 $\omega\in B\}$，如图 1-3 中阴影部分所示。

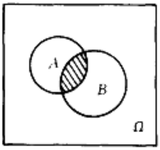

图 1－2

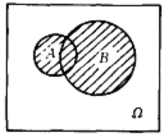

图 1－3

由于组成 $A\cup B$ 的样本点必至少来自 A,B 之一(可能来自两者)。反之出现在 A 或者 B 中的样本点必出现在 $A\cup B$ 中,所以事件 $A\cup B$ 发生$\Leftrightarrow A,B$ 中至少有一个发生。

类似地,称 $\bigcup\limits_{i=1}^{n}A_i$ 为 n 个事件 $A_1,A_2,\cdots,A_n$ 的和事件;称 $\bigcup\limits_{i=1}^{\infty}A_i$ 为可列个事件 $A_1,A_2,\cdots$的和事件。

4° "事件 A 发生而 B 不发生"这一事件称为事件 A 与 B 的差事件,记为 $A-B=\{\omega\mid\omega\in A$,且 $\omega\notin B\}$,如图 1－4 中阴影部分所示。

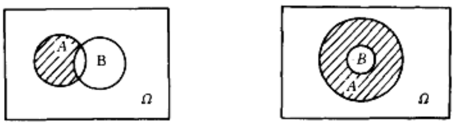

图 1－4

由于组成 $A-B$ 的样本点一定是在 A 中而不在 B 中。反之,出现在 A 中而不在 B 中的样本点必然出现在 $A-B$ 中,所以 $A-B$ 发生$\Leftrightarrow A$ 发生,且 B 不发生。

5° 若事件 A 与事件 B 不能同时发生,即 $AB=\varnothing$,则称事件 A 与 B 互不相容(或互斥),如图 1－5 所示。基本事件是两两互不相容的。

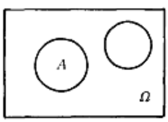

图 1－5

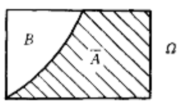

图 1－6

6° 若 A 是一个事件,令 $\bar{A}=\Omega-A$,称 $\bar{A}$ 是 A 的对立事件或逆事件。我们有 $A\cup\bar{A}=\Omega$, $A\bar{A}=\varnothing$。所以,一次试验中,A 和 $\bar{A}$ 必有且仅有一个发生,如图 1－6 所示。

例 1.5 在例 1.2 中 设 $A=\{1,3\}$, $B=\{0,3\}$,则

$A\cup B=\{0,1,3\}$；

$A\cap B=\{3\}$；

$A-B=\{1\}$；

$\bar{A}=\{0,2\}$；

$\overline{B-\bar{A}}=\overline{\{3\}}=\{0,1,2\}$。

这里大家会发现，事件间的关系及运算与集合间的关系及运算之间是完全可以互相类比的。对初学概率论的读者来说，重要的是要学会用概率论的语言来解释集合间的关系及运算，并能运用它们。

例 1.6 设 A,B,C 是 Ω 中的随机事件，则事件

"A 与 B 发生，C 不发生"可表示成 $AB\bar{C}$；

"A,B,C 中至少有两个发生"可表示成 $AB\cup AC\cup BC$；

"A,B,C 中恰好发生两个"可表示成 $AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$；

"A,B,C 中有不多于一个事件发生"可表示成 $\bar{A}\,\bar{B}\cup\bar{A}\,\bar{C}\cup\bar{B}\,\bar{C}$。

事件的运算满足下述规则。设 $A,B,C,A_i(i=1,2,\cdots,n)$ 为事件，则有

(1)**交换律** $A\cup B=B\cup A,AB=BA$

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

$(AB)C=A(BC)$

(3)**分配律** $(A\cup B)\cap C=AC\cup BC$

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

(4)**德摩根(De Morgan)律**

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

一般地　$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i}$；　$\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}$

我们仅证明分配律第二式，其余请读者完成。

证 设事件$(A\cap B)\cup C$发生，则 $A\cap B$ 及 C 必至少发生其一。

若 $A\cap B$ 发生，则必有 A,B 同时发生，故有 $A\cup C$ 与 $B\cup C$ 同时发生，即$(A\cup C)\cap(B\cup C)$发生。

若 C 发生，则必有 $A\cup C$ 与 $B\cup C$ 同时发生，亦有$(A\cup C)\cap(B\cup C)$发生。故有

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

反之，若$(A\cup C)\cap(B\cup C)$发生，则必有 $A\cup C$ 及 $B\cup C$ 同时发生。$A\cup C$ 发生，有 A，C 至少发生其一。$B\cup C$ 发生有 B,C 至少发生其一。

若 A,C 中 C 不发生，则 A 一定发生，且 B,C 中 B 也一定发生(C 不发生)，则 $A\cap B$ 一定发生，则有$(A\cap B)\cup C$ 必发生。

若 A,C 中 C 发生(不论 A,B 是否发生)，同样有$(A\cap B)\cup C$ 必发生。故有

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

故

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

上述事件运算规则也可以用集合论的方法加以证明，读者可以把刚才用概率论语言的证明"翻译"成集合论的语言，并证明其余各式以为练习。

§1.3 频率与概率

一、频率

前面我们看到随机试验有多个试验结果，每次试验之前不能确定这次试验出现哪个结果。但我们可以问：在一次试验中某些事件发生的可能性有多大？例如为了确定水坝的高度，就要知道河流在造水坝地段"每年最大洪水达到某一高度"这一事件发生的可能性大小。对于一个随机事件来说，它发生可能性大小的度量是由它自身决定的，并且是客观存在的，是随机事件自身的一个属性。我们希望将这一可能性量化，用一个适当的数表示出来。为此我们首先引入频率的概念，它是事件发生可能性大小的度量——概率的一个近似。

在相同条件下，重复进行了 n 次试验，在这 n 次试验中，事件 A 发生的次数 n_A 称为事件 A 发生的频数，比值 n_A/n 称为事件 A 发生的频率，记为 $f_n(A)$。

频率与概率的关系可以通过下面大家熟悉的问题加以说明。在实际测量长度和面积时，对于给定的一根棒，它有自身的"客观"的长度。长度的得来，我们可以用尺或仪器去测量，不论尺或仪器多么精确，测得的数据总是稳定在木棒"真实长度"值的附近。事实上，我们也常把测量所得的值当作真实的"长度"。我们的概率就好比这个"真实长度"，频率就好比这个"测量值"。这个类比不仅帮助我们去理解概率和频率之间的内在关系，也启示了更深刻的事实：概率与长度、面积等量一样，应该具有"可度量"的性质。

事件 A 发生的频率 $f_n(A)=n_A/n$ 具有下述性质：

(1)非负性：$f_n(A)\geqslant 0$；

(2)规范性：$f_n(\Omega)=1$；

(3)有限可加性：若 $A_1,A_2,\cdots,A_k$ 是一列两两互不相容的事件，则

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

证(3) 若 $A_1\cup A_2\cup\cdots\cup A_k$ 发生，意味着 $A_1,A_2,\cdots,A_k$ 中至少发生其中之一，又因为 $A_1,A_2,\cdots,A_k$ 互不相容(即任何两两不能同时发生)，所以 $A_1\cup A_2\cup\cdots\cup A_k$ 发生的次数一定是 $A_1,A_2,\cdots,A_k$ 分别发生次数之和，即 $n_{A_1\cup A_2\cdots\cup A_k}=n_{A_1}+n_{A_2}+\cdots+n_{A_k}$。因此有

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

从频率的定义我们不难发现：频率具有随机波动性。因为每次试验中事件 A 的发生与否具有不确定性，所以一般情况下，对于同样的 n，n_A 不一定相同，因而所得的 $f_n(A)$不尽相同。如前所述，事件发生的概率应该是唯一的，因此如果我们将概率代以频率显然是不合适的。但是前面分析过频率的本质就是概率，因而我们有理由要求频率的上述性质也是概率所应该具有的。由此得到启发，我们可以给出如下度量事件发生可能性大小的概率的确切定义。

二、概率

设 E 是随机试验，Ω 是其样本空间，对每一个随机事件 A 都有一个实数$P(A)$与之对应，称 $P(A)$为事件 A 的概率，如果集合函数 $P(\cdot)$满足下列条件：

(1)非负性：对于任一事件 A，有 $P(A)\geqslant 0$；

(2)规范性：$P(\Omega)=1$；

(3)可列可加性：设 $A_1,A_2,\cdots$是一列两两互不相容的事件，即对于 $i\neq j$，$A_iA_j=\varnothing(i,j=1,2,\cdots)$，有

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots \tag{1.1}$$

第五章中将证明，当 $n \to \infty$ 时频率 $f_n(A)$ 在一定意义下趋近于概率 $P(A)$，也将从理论上再一次表明频率的本质就是概率。

由概率的定义，可以导出概率的一些重要性质。

性质 1 $P(\varnothing) = 0$

证 令 $A_n = \varnothing (n = 1,2,\cdots)$，则 $\bigcup\limits_{n=1}^{\infty} A_n = \varnothing$，且 $A_iA_j = \varnothing, i \neq j$。由概率的可列可加性 (1.1) 得

$$P(\varnothing) = P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(\varnothing)$$

而实数 $P(\varnothing) \geqslant 0$，故由上式得 $P(\varnothing) = 0$。

性质 2 若 $A_1, A_2, \cdots, A_n$ 是两两互不相容的事件，则有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{1.2}$$

(1.2)式称为概率的有限可加性。

证 令 $A_{n+1} = A_{n+2} = \cdots = \varnothing$，即有 $A_iA_j = \varnothing, i \neq j, i,j = 1,2,\cdots$。由(1.1)式得

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) &= P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{n} P(A_k) + 0 \\ &= P(A_1) + P(A_2) + \cdots + P(A_n) \end{aligned}$$

性质 3 设 A, B 是两个事件，若 $A \subset B$，则有

$$P(B - A) = P(B) - P(A) \tag{1.3}$$

$$P(B) \geqslant P(A) \tag{1.4}$$

证 由 $A \subset B$ 知 $B = A \cup (B - A)$（参见图 1-1），且 $A(B - A) = \varnothing$，再由概率的有限可加性(1.2)，得

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

(1.3)得证；又由概率定义中的(1)，$P(B - A) \geqslant 0$，故

$$P(B) \geqslant P(A)$$

性质 4 对于任一事件 A，

$P(A) \leqslant 1$

证 因 $A \subset \Omega$，由性质 3 得

$P(A) \leqslant P(\Omega) = 1$

性质 5 对于任一事件 A，有

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

证 因 $A \cup \bar{A} = \Omega$，且 $A\bar{A} = \varnothing$，由(1.2)式，得

$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$

性质 6 对于任意两事件 A, B 有

$$P(A \cup B) = P(A) + P(B) - P(AB) \tag{1.5}$$

证 因 $A \cup B = A \cup (B - AB)$（参见图 1-3），且 $A(B - AB) = \varnothing$，$AB \subset B$，故由(1.2)及(1.3)得

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

(1.5)式还能推广到多个事件的情况。一般地，对于任意 n 个事件 $A_1, A_2, \cdots, A_n$，可以用归纳法证得

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_iA_j)$$
$$+ \sum_{1 \leq i < j < k \leq n} P(A_iA_jA_k) + \cdots + (-1)^{n-1} P(A_1A_2\cdots A_n) \tag{1.6}$$

§1.4 古典概型

前面已经提到，一个随机试验在数学上用样本空间 Ω、随机事件及其概率 $P(\cdot)$ 来描述。对于一个随机事件 A，如何寻求它的概率 $P(A)$ 是概率论的一个基本的课题。我们先讨论一类最简单的随机试验，它具有下述待征：

(1)样本空间的元素(样本点)只有有限个。不妨设为 n 个，并记作 $\omega_1, \omega_2, \cdots, \omega_n$；

(2)每个基本事件发生的可能性相同，即 $P(\omega_1) = P(\omega_2) = \cdots = P(\omega_n)$。

这类试验称为等可能概型，又称古典概型。它在概率论发展初期曾是主要的研究对象，在概率论中有很重要的地位。一方面，因为它比较简单，许多概念既直观而又容易理解；另一方面，它又概括了许多实际问题，有着广泛的应用。

下面我们来研究古典概型中事件概率的计算，其样本空间 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$。由于基本事件是两两互不相容的，于是由概率的有限可加性有

$$1 = P(\Omega) = P(\omega_1 \cup \omega_2 \cup \cdots \cup \omega_n)$$
$$= P(\omega_1) + P(\omega_2) + \cdots + P(\omega_n) = nP(\omega_i), \quad i = 1, 2, \cdots, n$$

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

对任一随机事件 A，如果 A 包含 k 个基本事件，不妨设 $A = \omega_{i_1} \cup \omega_{i_2} \cup \cdots \cup \omega_{i_k}$，$(1 \leq i_1 < i_2 < \cdots < i_k \leq n)$。则有

$$P(A) = \sum_{j=1}^{k} P(\omega_{i_j}) = \frac{k}{n} = \frac{A\text{ 中所含的基本事件数}}{\Omega\text{ 中的基本事件总数}}$$
$$= \frac{A\text{ 的有利事件数}}{\Omega\text{ 中的基本事件总数}}$$

A 中所含的基本事件数，习惯上常称为 A 的有利事件数。不难验证上述的概率 $P(\cdot)$ 的确具有非负性、规范性和有限可加性。

例 1.7 书架上有 5 本中文书、3 本英文书、2 本日文书，求从中任取一本时得到的是外文书的概率。

解 $\Omega = \{5$ 本中文书，3 本英文书，2 本日文书(即全部 10 本书)$\}$，记 $A = \{3$ 本英文书$\}$，$B = \{2$ 本日文书$\}$，则

$$P(A) = \frac{A\text{ 中所含的基本事件数}}{\Omega\text{ 中基本事件总数}} = \frac{3}{10}$$
$$P(B) = \frac{B\text{ 中所含的基本事件数}}{\Omega\text{ 中基本事件总数}} = \frac{2}{10}$$

由于取到的一本书不会既是英文书，又是日文书，所以 $P(AB) = 0$，则

$$P(A \cup B) = P(A) + P(B) - P(AB) = \frac{1}{2}$$

当一个古典概型试验的样本空间中所含样本点数较多时，我们就很难(或事实上不可能)把样本点一一列出，然后进行数数。这时要用到一些计数工具，例如排列和组合的知识。

即使是样本点不多的情形,有时运用排列和组合的知识求古典概型中事件发生的概率,也比直接数数来得简单。样本点数较多时,只需分别写出 Ω 和 A 中的样本点的个数(即基本事件数)进而计算概率。有一点读者必须注意:为求一个事件的概率,样本空间可以有不同的取法,对结果的计算难易上也会有所差别。同时一定要认清,基本事件总数和有利事件数的计算都要在同一个样本空间中进行,否则会引起混淆,并导致谬误。此外还有一些需要注意的问题,我们在下面的例题中将看到。

例 1.8 一口袋中有 5 只白球 4 只黑球,今从袋中任取两球,问取得白球黑球各一个的概率是多少?

解法一 袋中共有 9 只球,从这 9 只球中任取两只(不论次序)共有 $C_9^2=\dfrac{9\times 8}{2}=36$ 种取法,所以样本空间由 36 个样本点组成,这些样本点出现的机会是彼此相同的。

用 A 表示事件"取得白球黑球各一个",它是在 5 个白球中任取其一,4 个黑球中任取其一,因此由 $5\times 4=20$ 个样本点组成,故

$$P(A)=\frac{20}{36}=\frac{5}{9}$$

解法二 从袋中任取两球,我们可以看作是先取一个球(不放回袋中),再取一个球,这时一切可能的取法总数为 $9\times 8=72$,它们出现的可能性也是相同的。这时取得白球黑球各一个也可以区分为"先取出一个白球,后取出一个黑球"(有 $5\times 4=20$ 种取法)和"先取出一个黑球,后取出一个白球"(有 $4\times 5=20$ 种取法),这时组成事件 A 的样本点为 40,因此

$$P(A)=\frac{40}{72}=\frac{5}{9}$$

这个例子说明,在有些问题中样本空间有不同的选择,但无论怎样选择,样本空间的各个样本点应具有等可能性,并且在选定样本空间后,我们所关心的事件能表示成所选样本空间的子集,同时样本空间中样本点与组成所关心事件的样本点必须是在同一试验方式下。例如在解法一中样本空间中的点由两个球组成,不论次序。如果把 5 个白球标以白$_1$、白$_2$、白$_3$、白$_4$、白$_5$,4 个黑球标以黑$_1$、黑$_2$、黑$_3$、黑$_4$,那么(白$_1$,黑$_2$)与(黑$_2$,白$_1$)是同一个样本点。因而在考虑事件 A 由哪些样本点组成时,也只顾及样本点由哪两个球组成,不计次序。但在解法二中,(白$_1$,黑$_2$)与(黑$_2$,白$_1$)视为两个不同的样本点,在考虑事件 A 由哪些样本点组成时,不仅要考虑由哪两个球组成,还要考虑两个球取出的次序,必须与样本空间的选择一致起来。

例 1.9 设有 6 件产品,其中 4 件合格、2 件不合格(假设外观上没有差别)。今从中取两次,每次随机地取一件。考虑两种取法:(a)第一次取一件产品检验后放回,与其他产品充分混放后再取一件,这种取法称为放回抽样;(b)第一次取一产品检验后不放回,第二次从剩余的产品中再取一件,这种取法称为不放回抽样。试分别就上面两种情况求:(i)取到的两件产品都是合格品的概率;(ii)取到的两件产品中至少有一件是合格品的概率。

解 (a)放回抽样的情况

以 A、B、C 分别表示事件"取到的两件产品都是合格品"、"取到的两件产品都是次品","取到的两件产品中至少有一件是合格品",这里 $C=\bar{B}$。

依次取两件产品,每一种取法为一个基本事件,此时样本空间中仅包含有限个元素,且由对称性知每个基本事件发生的可能性相同,属于古典概型。

第一次取产品共有 6 件可取,第二次也有 6 件可取,由组合法的乘法原理,共有 6×6 种

取法。即样本空间中元素总数为 6×6。对于事件 A 而言，由于第一次有 4 件合格品可供抽取，第二次也有 4 件可供抽取，由乘法原理共有 4×4 种取法，即 A 中包含 4×4 个元素。同理，B 中包含 2×2 个元素。于是

$$P(A)=\frac{4\times4}{6\times6}=0.444$$

$$P(B)=\frac{2\times2}{6\times6}=0.111$$

$$P(C)=P(\bar{B})=1-P(B)=0.889$$

(b)不放回抽样的情况

由读者自己完成。

例 1.10 设有 n 个人，每个人都等可能地被分配到 N 个房间中的任意一间去住($n\leqslant N$)，求下列事件的概率。

(1)指定的 n 个房间各有一个人住；

(2)恰好有 n 个房间，其中各住一个人。

解 (1)因为每一个人有 N 个房间可供选择，所以 n 个人住的方式共有 N^n 种，它们是等可能的，指定的 n 个房间各有一人住，其可能总数为 n 个人的排列 $n!$，于是

$$P_1=\frac{n!}{N^n}$$

(2) n 个房间可以在 N 个房间中任意选取，其总数有 C_N^n 个，对选定的 n 个房间，按前述的讨论可知有 $n!$ 种分配方式，所以恰有 n 个房间其中各住一个人的概率为

$$P_2=\frac{C_N^n\cdot n!}{N^n}=\frac{N!}{N^n(N-n)!}$$

这个例子常称为“分房问题”。有许多问题和本例具有相同的数学模型，如把本例中的“人”理解为“粒子”，“房间”理解为粒子所处的能级，则“分房问题”所描述的模型就是统计物理学中的马克斯威尔-波尔兹曼(Maxwell-Boltzmann)统计。如果 n 个人(即“粒子”)是可分辨的，则上述模型即对应于玻色－爱因斯坦(Bose-Einstein)统计(见本章习题 13)。

由以上的例题我们看到，求解古典概型问题的关键是在寻找基本事件总数和有利事件数，但正面求这两个数并不都是很容易的，有时要讲究一些技巧。要掌握这些技巧，当然需要一些艰苦的训练。

例 1.11 某班级有 n 个人($n\leqslant365$)，问至少有两人的生日在同一天的概率为多大？

解 假定按 365 天计算，把 365 天当作 365 个“房间”，那么问题就可以归结为上例。这时“n 个人的生日全不相同”就相当于上例中的(2)。令 $A=\{n$ 个人中至少两个人的生日相同$\}$，则 $\bar{A}=\{n$ 个人的生日全不相同$\}$，由上例(2)知

$$P(\bar{A})=\frac{N!}{N^n\cdot(N-n)!}$$

而

$$P(A)+P(\bar{A})=1$$

于是

$$P(A)=1-\frac{N!}{N^n\cdot(N-n)!}\qquad(N=365)$$

在这个例子中，如果直接求 $P(A)$是比较麻烦的，而利用对立事件求解那就简便多了。

这个例子是历史上有名的“生日问题”，对不同的一些 n 值，计算的 $P(A)$值如下表：

n	10	20	23	30	40	50
$P(A)$	0.12	0.41	0.51	0.71	0.89	0.97

上表所列的答案是足以引起多数读者惊奇的,因为"一个班级中至少有两个人的生日相同"这件事情发生的概率,并不如大多数人直觉中想象的那样小,而是相当大。由表中可以看出,当班级中的人数为23时就有半数以上的班级会发生这件事情,而当班级的人数达到50时,竟有97%的班级会发生上述事件。当然,这里讲的"半数以上"、"有97%"都是就概率而言。这个例子告诉了我们,"直觉"并不很可靠,这就有力地说明了研究随机现象统计规律的重要性。

例 1.12　某接待站在某一周曾接待过12次来访,已知所有这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中去接待站是等可能的,则12次接待来访者都在周二、周四的概率为 $2^{12}/7^{12}=0.0000003$,即千万分之三。人们在长期的实践中总结得到"概率很小的事件在一次试验中实际上几乎是不发生的"(称之为实际推断原理)。现在概率很小(只有千万分之三)的事件在一次试验中竟然发生了,因此有理由怀疑假定的正确性,从而推断接待站不是每天都接待来访者,即认为其接待时间是有规定的。

§1.5　几何概型

在概率论发展的早期,人们就已经注意到只考虑那种仅有有限个等可能结果的随机试验是不够的,还必须考虑试验结果是无穷多个的情形。这中间最简单的一类是试验结果有无穷多个而又有某种等可能性的场合,本节例1.14就是著名的一例。

如果我们在一个面积为 S_Ω 的区域 Ω 中,等可能地任意投点(见图1-7),这里"等可能"的确切意义是这样的:设在区域 Ω 中有任意一个小区域 A,若它的面积为 S_A,则点落入 A 中的可能性大小与 S_A 成正比,而与 A 的位置及形状无关。如果"点落入小区域 A"这个随机事件仍记作 A,则由 $P(\Omega)=1$,有 $P(A)=S_A/S_\Omega$。一般地,我们有如下的定义:

定义 1.1　(1)进行一次试验相当于向一个几何体(有限线段、平面区域、曲面、空间区域或它们的并)Ω 中取一点 g,Ω 有非零的有限测度(长度、面积或体积);(2)对 Ω 内任何一个小区域 A,事件"点取自 A"的概率与 A 的测度成正比,与 A 的形状及在 Ω 中的位置无关。称满足(1)、(2)的随机试验为几何概型。

几何概型中,事件"点取自 A"的概率定义为:

$$P(A)=\frac{\mu(A)}{\mu(\Omega)}$$

其中 $\mu(A)$ 表示几何体 A 的测度,如此定义的概率称为几何概率(参见图1-8)。

下面我们看几个例子,注意一下实际问题中是如何把随机试验归结为向一个几何体取点的问题。

例 1.13(会面问题)　甲、乙两人约定在6时到7时之间在某处会面,并约定先到者应等候另一个人一刻钟,过时即可离去,求两人能会面的概率。

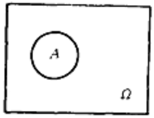

图 1-7

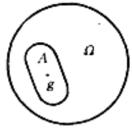

图 1-8

解 以 x 和 y 分别表示甲、乙两人到达约会地点的时刻，每次试验 (x,y) 相当于在边长为 60 的正方形区域

$$\Omega = \{(x,y) \quad : \quad 0 \leqslant x, y \leqslant 60\}$$

中取一点。而两人能够会面的充要条件是 $|x-y| \leqslant 15$，如图 1-9。

由于甲、乙两人在 $[0,60]$ 时间区间中任何时刻到达是等可能的，而且两人互不影响，所以 (x,y) 在 Ω 的任何可求积子区域内取点的可能性只与区域的面积大小成正比而与其形状、位置无关，于是我们把会面问题转化成向平面区域 Ω 随机取点的问题。由等可能性有：

$$P(A) = \frac{S_A}{S_\Omega} = \frac{60^2 - 45^2}{60^2} = \frac{7}{16}$$

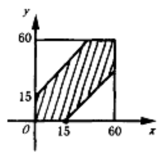

图 1-9

例 1.14(投针问题) 平面上有一簇平行线，其中任何相邻的两线距离都是 $a(a>0)$。向平面任意投一长为 $l(l<a)$ 的针，试求针与一条平行线相交的概率。

解 设 x 是针的中点 M 到最近的平行线的距离，φ 是针与此平行线的交角。于是投针问题就相当于向平面区域

$$\Omega = \{(\varphi, x) : 0 \leqslant \varphi \leqslant \pi, 0 \leqslant x \leqslant \frac{a}{2}\}$$

取点的几何概型。

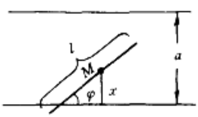

图 1-10

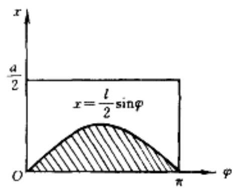

图 1-11

针与平行线之一相交的充分必要条件是 (φ, x) 满足条件

$$0 \leqslant x \leqslant \frac{l}{2}\sin\varphi, 0 \leqslant \varphi \leqslant \pi$$

即随机点取自平面区域

$$A = \{(\varphi, x) : 0 \leqslant x \leqslant \frac{l}{2}\sin\varphi, 0 \leqslant \varphi \leqslant \pi\}$$