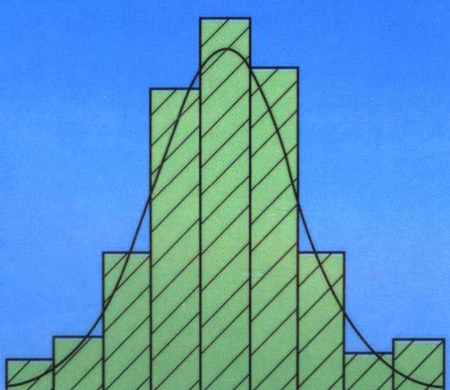
概率论与数理统计

（第一版）

顾玉娣　杨纪龙　编



航空工业出版社

2002

概率论与数理统计

（第一版）

顾玉娣　杨纪龙　编

航空工业出版社

2002

内 容 提 要

本书详细介绍了概率论与数理统计的基本概念、基本理论及其应用。内容包括概率空间、随机变量与分布函数、随机向量及其分布、数字特征和特征函数、极限定理、抽样分布、参数估计、假设检验、回归分析、方差分析与正交设计。本书既可作为教材供非数学类研究生及数学类学生使用，也可供工程技术人员自学与参考

图书在版编目(CIP)数据

概率论与数理统计/顾玉娣等编.—北京：航空工业出版社，2002
ISBN 7-80134-978-4

Ⅰ.概… Ⅱ.顾… Ⅲ.①概率论—高等学校—教材②数理统计—高等学校—教材 Ⅳ.021

中国版本图书馆 CIP 数据核字（2002）第 023694 号

航空工业出版社出版发行
（北京市安定门外小关东里 14 号 100029）

南京航空航天大学印刷厂印刷　　全国各地新华书店经售
2002 年 6 月第 1 版　　2002 年 6 月第 1 次印刷
开本 787×1092 1/16　　印张 14 5 字数 370 千字
印数 1—1300　　定价 25 00 元

前　　言

本教材是作者在为非数学类研究生编写的《概率论与数理统计》讲义基础上修改而成的。本书可供非数学类研究生使用，并兼顾工科院校数学本科生使用。

全书共分十章，前五章是概率论部分，主要讲述概率论的基本概念和基本结论，由顾玉娣编写；后五章是数理统计部分，主要讲述数理统计基本概念和常用统计方法，由杨纪龙编写，全书由东南大学朱道元教授审阅。

本书的出版得到了南京航空航天大学研究生院、教务处及教材科的领导和同志们的支持和帮助，在此表示衷心的感谢！

由于编者水平有限，书中的缺点和错误在所难免，恳请同行专家和读者批评指正。

编　者

2002.3

目　录

第一章　概率空间

第一节　随机事件及其运算

一、随机现象和随机试验

在自然界里，在生产实践和科学实验中，人们观察到的各种各样的自然现象和社会现象可归结为两类：一类是**必然现象**或确定性现象，即在一定的条件下必然会出现的现象，例如在没有外力作用的条件下，作等速直线运动的物体必然继续作等速直线运动。又如在标准大气压下，水加热到 100 ℃时必然会沸腾等。另一类是**随机现象**或偶然现象，即在完全相同的条件下重复试验，每次结果未必相同。例如掷一枚均匀硬币，可能出现两种不同的结果，正面向上或反面向上，但究竟出现哪种结果事先无法确切知道。又如进入某商店购物的顾客数，可能是任意一个非负整数，但事先无法预言每天顾客的确切数字。

是不是这些随机现象没有什么规律呢？事实并非如此。人们通过长期的反复观察和实验，逐渐发现随机现象也呈现某种规律性。例如多次重复地掷一枚均匀硬币，正面朝上的次数大致有一半，因此对于随机现象，虽然一次实验的结果呈现出不确定性，但在大量重复实验中其结果却具有一定的规律，我们称之为**统计规律性**。概率论与数理统计正是研究随机现象的统计规律性的数学分支。

我们把对自然现象进行观察或进行实验统称为试验。如果试验在相同的条件下可以重复进行，每次试验的结果不止一个且事先能明确该试验的所有可能的结果，而每次试验的结果事前不能预言，我们称之为**随机试验**，常用字母 E 表示。

二、样本空间与随机事件

1. 样本空间

为了研究随机试验，首先需要知道这个试验可能出现的所有结果，每一个可能的结果称为随机试验 E 的一个**样本点**，一般用 ω 表示。由 E 的所有样本点组成的集合称为 E 的**样本空间**，用 Ω 表示。

例 1.1　掷一枚均匀硬币，观察其出现正、反面的情况，试验的结果只有两种可能：正面和反面。记 ω_1 为正面，ω_2 为反面，则样本空间 $\Omega=\{\omega_1,\omega_2\}$。

例 1.2　同时掷两枚均匀硬币，观察出现正、反面的情况，其样本点为 $\omega_1=$（正，正），$\omega_2=$（正，反），$\omega_3=$（反，正），$\omega_4=$（反，反），则样本空间 $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$。

例 1.3　观察某商店每天的顾客数，则样本点为 $\omega_0,\omega_1,\omega_2,\cdots$，其中 $\omega_k=$（有 k 个顾客），$k=0,1,2,\cdots$，样本空间 $\Omega=\{\omega_k: k=0,1,2,\cdots\}$。

例 1 4　观察一批灯泡的寿命，样本空间 $\Omega=\{t\,|\,t\geqslant 0\}$。

样本空间可以分为：

(1) **有限样本空间**：只有有限个样本点的样本空间，如例 1.1，例 1.2 中的样本空间。

(2) **可列样本空间**:有无穷多个样本点,但这些样本点可以依照某种次序排列出来,我们称它的样本点数为可列无穷多个,如例 1.3 中的样本空间。

(3) **无穷样本空间**:有无穷多个样本点,它们充满一个或几个区间(区域),不是一个可列集,如例 1.4 中的样本空间。

有限样本空间及可列样本空间统称为**离散样本空间**。

2. 随机事件

样本点的某个集合称为**随机事件**,简称为**事件**,常用字母 A,B,C 等表示,由单个样本点构成的集合称为**基本事件**,由多个样本点构成的集合称为**复合事件**。在一次试验中,当且仅当某一事件所包含的某一个样本点出现时,称该事件发生。

我们把样本空间 Ω 也作为一个事件,由于在每一次试验中必然出现 Ω 中的某个样本点,所以 Ω 必然发生,称 Ω 为**必然事件**。我们把空集 $\varnothing$ 也作为一个事件,在每次试验中它都不会发生,称 $\varnothing$ 为**不可能事件**。

如例 1.2 中,事件 $A=\{\omega_1\}$ 表示事件"两次均为正面",事件 $B=\{\omega_1,\omega_2,\omega_3\}$ 表示事件"至少出现 1 次正面"。

三、事件的关系及运算

概率论的重要内容之一是希望从简单事件的概率推算出复杂事件的概率。分析与研究事件之间的关系及运算不仅帮助我们更深刻地认识事件的本质,而且可以大大简化一些复杂事件的概率计算。

1. 事件的包含与相等

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,记为 $B\supset A$,或 $A\subset B$,可由图 1-1(1)表示。若 $A\supset B$ 且 $B\supset A$,则称事件 A 与 B 相等,记作 $A=B$。

2. 事件的和(并)

"事件 A 与事件 B 至少有一个发生"这一事件称为事件 A 与 B 的和(或并)事件,记作 $A\cup B$,即 $A\cup B=\{\omega:\omega\in A$ 或 $\omega\in B\}$,如图 1-1(2)。

3. 事件的积(交)

"事件 A 与事件 B 同时发生"这一事件称为事件 A 与 B 的积(或交)事件,记作 $A\cap B$ 或 AB,即 $A\cap B=\{\omega:\omega\in A$ 且 $\omega\in B\}$,如图 1-1(3)。

显然对任一事件 A,有 $A\cap\Omega=A$,$A\cap\varnothing=\varnothing$。

4. 事件的差

"事件 A 发生而事件 B 不发生"这一事件,称为事件 A 与 B 的差事件,记为 $A-B$,即 $A-B=\{\omega:\omega\in A$ 且 $\omega\notin B\}$,如图 1-1(4)。

5. 事件的互不相容(互斥)

若 $A\cap B=\varnothing$,则称事件 A 与 B 互不相容(或互斥),即事件 A 与 B 不能同时发生,如图 1-1(5)。基本事件是互不相容的。

6. 逆事件(对立事件)

"事件 A 不发生"这一事件,称为事件 A 的逆事件(或对立事件),记作 $\overline{A}$,即 $\overline{A}=\{\omega:\omega\notin A\}$,也可表示为 $\overline{A}=\Omega-A$,如图 1-1(6)。

若事件 A,B 满足条件:$A\cup B=\Omega$,$A\cap B=\varnothing$,则称 B 是 A 的逆事件(或对立事件)记作

$B=\overline{A}$，也可称 A 是 B 的逆事件，记作 $A=\overline{B}$，因此称 A 与 B 互为逆事件，显然有 $A-B=A\overline{B}$。

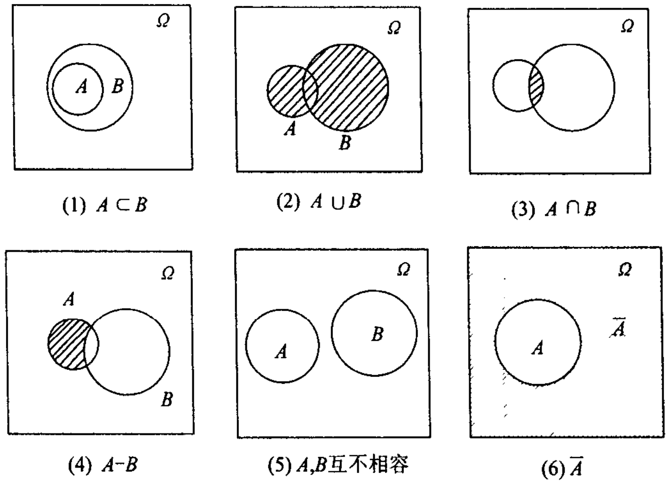


图 1-1　事件的运算

不难把事件的和与事件的积推广到有限多个事件及可列多个事件的情形。例如，对于 n 个事件 $A_1,A_2,\cdots,A_n$，用 $\bigcup\limits_{i=1}^{n}A_i$ 表示 $A_1,A_2,\cdots,A_n$ 中至少有一个发生，称为 $A_1,A_2,\cdots,A_n$ 的和，用 $\bigcap\limits_{i=1}^{n}A_i$ 表示 $A_1,A_2,\cdots,A_n$ 同时发生，称为 $A_1,A_2,\cdots,A_n$ 的积。

对于可列个事件，我们定义

$$\bigcup_{i=1}^{\infty}A_i=\lim_{n\to\infty}\bigcup_{i=1}^{n}A_i,\quad \bigcap_{i=1}^{\infty}A_i=\lim_{n\to\infty}\bigcap_{i=1}^{n}A_i$$

由于事件是样本点的某个集合，即事件是样本空间的某个子集，所以事件间的关系及运算与集合论中的关系及运算是完全相似的，仅仅是在叙述中的某些术语不同而已。不过，我们还是要学会用概率论的语言来解释这些关系及运算。

四、事件运算的简单性质

设 A,B,C 是任意事件，那么它们满足：

(1) 交换律：$A\cup B=B\cup A,A\cap B=B\cap A$　(1.1)

(2) 结合律：$(A\cup B)\cup C=A\cup(B\cup C)$　(1.2)

$(A\cap B)\cap C=A\cap(B\cap C)$　(1.3)

(3) 分配律：$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$　(1.4)

$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$　(1.5)

(4) 对偶律(De Morgan 公式)

$\overline{A\cup B}=\overline{A}\cap\overline{B};\overline{A\cap B}=\overline{A}\cup\overline{B}$　(1.6)

对于 n 个事件，甚至对于可列个事件，性质(1)，(2)，(4)也成立。例如，设 $A_1,A_2,\cdots$ 为可列个事件，则有

$$\overline{\bigcup_{i=1}^{\infty}A_i}=\bigcap_{i=1}^{\infty}\overline{A_i},\ \overline{\bigcap_{i=1}^{\infty}A_i}=\bigcup_{i=1}^{\infty}\overline{A_i}\tag{1.7}$$

例 1.5 设 A,B,C 是同一随机试验的三个事件，试用 A,B,C 及其运算表示下列事件：

(1) A,B 都发生而 C 不发生；

(2) A,B,C 中恰有两个事件发生；

(3) A,B,C 中至多有两个事件发生；

(4) A,B,C 中至少有两个事件不发生。

解 (1) $AB\overline{C}$

(2) $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$

(3) $\overline{A}\ \overline{B}\overline{C}\cup A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\ \overline{B}C\cup AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC=\overline{ABC}=\overline{A}\cup\overline{B}\cup C$

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

第二节　概率的定义及计算

研究随机现象，不仅要知道可能出现哪些事件，更重要的是要弄清各事件发生的可能性的大小，以揭示出这些事件的内在统计规律。事件在一次试验中发生的可能性大小这一刻划事件的数量指标通常称为该事件的**概率**。本节中，我们提出在一些简单情形下合理地规定并计算随机事件的概率的方法，然后再从理论上完善概率的定义，引进概率的公理化定义。

一、频率与概率的统计定义

1. 频率及其稳定性

在实际问题中，我们发现，当考虑某个事件 A 发生的可能性的大小时只要在相同的条件下作大量的重复试验，事件 A 发生的频率呈现某种稳定现象，即当试验次数增加时，事件 A 发生的频率稳定于某一常数的附近，而且偏离的可能性很小，这一现象称为频率的稳定性。频率的这一性质清楚地说明刻划事件发生的可能性大小的数——概率是客观存在的，而且它启发我们当试验次数足够大时可用事件的频率作为该事件的概率的一种度量，概率的这种定义称为概率的统计定义。

定义 1.1 设在相同的条件下进行了 n 次试验，事件 A 发生了 n_A 次，则称

$$f(A)=\frac{n_A}{n} \tag{1.8}$$

为事件 A 在这 n 次试验中的频率。

频率具有以下性质：

(1) 对任何事件 A，$0\leqslant f_n(A)\leqslant 1$；

(2) $f_n(\Omega)=1$；

(3) 若事件 $A_1,A_2,\cdots,A_k$ 为两两互不相容的事件，则

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

2. 概率的统计定义

定义 1.2 设在相同条件下重复试验了 n 次，事件 A 发生的频率为 $f_n(A)$，当 n 足够大时我们将 $f_n(A)$ 作为事件 A 的概率的一个度量，称为事件 A 的**统计概率**。

显然统计概率具有频率的性质。

统计概率直观、具体，但在理论上有明显的缺点。统计概率定义的依据是“频率的稳定性”，然而这一事实不能用通常数学中的数列极限来描述(而只能用随机变量序列的收敛性来描述，详见第五章第二节)，我们并不能认为试验次数为 $n+1$ 时的频率总会比试验次数为 n 时的频率更逼近所求的概率。因此，用频率求得的概率是很“粗糙”的。但是，随着电子计算机的科学计算技术的深入发展，用计算机随机模拟并计算事件的统计概率已在实际中得到广泛应用。

二、古典概率

下面我们研究一类简单而又特殊的随机试验，它具有两个特点：

(1) 随机试验 E 的全部可能结果只有有限多个，E 的样本空间记为 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$；

(2) 基本事件 $\{\omega_1\},\{\omega_2\},\cdots,\{\omega_n\}$ 出现的可能性相等。

这类随机试验 E 的数学模型称为**古典概型**。

定义 1.3 在古典概型中，对于任意事件 A，对应的概率 $P(A)$ 为

$$P(A)=\frac{\text{事件 } A \text{ 包含的基本事件数 } k}{\Omega \text{ 中基本事件总数 } n} \tag{1.9}$$

并称为**古典概率**。

等可能性是一种假设，实际应用中，在很多场合是由对称性(如掷硬币试验中认为硬币是均匀的)或某种均衡性(如摸球试验中认为球的大小相同)来判断基本事件的等可能性，并在此基础上计算事件的概率。

不难推出古典概率具有如下性质：

(1) 非负性：对任一事件 A，有 $P(A)\geqslant 0$；

(2) 规范性：$P(\Omega)=1$；

(3) 有限可加性：设事件 $A_1,A_2,\cdots,A_n$ 两两互不相容，则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$$

例 1.6 100 只同批生产的产品，按质量分类有 40 只属于一等品，60 只属于二等品，现从中任意抽取 3 次，每次 1 只，求其中有 2 只是一等品、1 只是二等品的概率。抽取方式分为两种：(1)放回抽样；(2)不放回抽样。

解 设 A 表示所求概率的事件

(1) 放回抽样

基本事件总数 $n=100^3$

事件 A 包含的基本事件数 $k=C_3^2\times 40^2\times 60$

则有 $P(A)=\dfrac{C_3^2\times 40^2\times 60}{100^3}=0.288$

注 上式可写为 $P(A)=C_3^2\times\left(\dfrac{40}{100}\right)^2\times\dfrac{60}{100}=C_3^2\times(0.4)^2\times 0.6$(参见本章第四节的独立试验概型)

(2) 不放回抽样

基本事件总数 $n=A_{100}^3$(将每只产品都看作是有区别的，因此用排列)

事件 A 包含的基本事件数 $k=C_3^2\times C_{60}^1\times A_{40}^2$

故有 $P(A)=\dfrac{C_3^2\times C_{60}^1\times A_{40}^2}{A_{100}^3}\approx 0.289$

注 上式也可以改写为

$$P(A)=\frac{C_3^2\times C_{60}^1\times A_{40}^2}{A_{100}^3}=\frac{C_{60}^1\times C_{40}^2}{C_{100}^3} \tag{1.10}$$

式(1.10)最右端分母 C_{100}^3 表示从 100 只产品中任取 3 只且一次同时取出,因此不考虑次序用组合计算。可见,不放回地逐一抽样 m 次,可以看作一次共抽 m 件,用组合计算更简便。

例 1.6(2)的更一般的提法为:一袋中装有 n 个球,其中 n_1 个带有号码"1",n_2 个带有号码"2",n_3 带有号码"3",$n_1+n_2+n_3=n$。从此袋中任取 $m(m\leqslant n)$ 个球,求恰有 m_i 个带有号码"i"($i=1,2,3$)的概率,其中 $m_1+m_2+m_3=m$,且 $m_i\leqslant n_i$,($i=1,2,3$)。

用式(1.10)相同的解法,得所求概率为

$$p=\frac{C_{n_1}^{m_1}C_{n_2}^{m_2}C_{n_3}^{m_3}}{C_n^m} \tag{1.11}$$

上式称为超几何概率,实际中的许多问题都可以看作这一模型。

从例 1.6 可以看出放回抽样与不放回抽样所得概率是不同的,但当被抽取的对象总数较大时,两种不同抽样方式所得概率相差较小。

例 1.7 袋中有 a 只黑球和 b 只白球,它们除颜色不同外,其它方面没有差别。现在把球随机地一只一只摸出来,求第 k 次摸出的一只是黑球的概率($1\leqslant k\leqslant a+b$)。

解法 1 把 a 只黑球和 b 只白球都看作是不同的(设想把它们编号),若把球摸出并依次排列,则可能的排法总数为$(a+b)!$,第 k 次摸出的是黑球有 a 种取法,而另外$(a+b-1)$次摸球相当于$(a+b-1)$只球进行全排列,有$(a+b-1)!$ 种不同的排法,故所求概率为

$$p=\frac{a\times(a+b-1)!}{(a+b)!}=\frac{a}{a+b}$$

解法 2 把 a 只黑球看作是没有区别的,把 b 只白球也看作是没有区别的,把球摸出并依次排列时,若把 a 只黑球的位置固定下来则其它位置必然是放白球。黑球的位置可以有 C_{a+b}^a 种选法,若第 k 次摸得黑球,则第 k 个位置必须放黑球,剩下的黑球可以放在 $a+b-1$ 个位置中任意 $a-1$ 个位置上,因此共有 C_{a+b-1}^{a-1} 种放法。所以所求概率为

$$p=\frac{C_{a+b-1}^{a-1}}{C_{a+b}^a}=\frac{a}{a+b}$$

两种解法的区别在于选取的样本空间不同,解法 1 中把球都看作是不同的,要考虑各黑球及各白球间的顺序,因此用排列;解法 2 中仅考虑球的不同颜色的区别而对同色球不加区别,因此用组合。

三、几何概率

以上古典概型的两个特点是"等可能性"和"有限性",但在实际问题中会遇到样本空间有无穷多个(不是可列个)样本点的情形。如何计算样本空间有无穷多个样本点但仍具有"等可能性"的事件的概率呢?

例如,某公共汽车站每隔 10 min 有一辆公共汽车停靠,某人到汽车站等候乘车,求他等待时间不超过 3 min 的概率。

由于公共汽车站每隔 10 min 有一辆汽车到达，我们自然认为某人到达车站的时间是在两辆汽车到站的时间之间，而且在这段时间之内的各个时刻是等可能的，只有当他到达车站的时间正好离第二辆汽车到达的时间不超过 3 min，则他的候车时间才不超过 3 min，因此相应的概率应该为$\frac{3}{10}=0.3$。

一般地，这类问题中，试验的所有可能结果可以用欧氏空间的某一区域 Ω 来表示，而"等可能性"的意义是：落在区域 Ω 中的任一可能出现的小区域 A 内的可能性都是可以度量的，且落在该小区域 A 内的概率与区域 A 的测度（如长度、面积、体积等）成正比而与 A 的位置及形状无关。

定义 1.4 设样本空间 Ω 的测度为 $\mu(\Omega)$，事件 A 的测度为 $\mu(A)$，则事件 A 的概率为

$$P(A)=\frac{\mu(A)}{\mu(\Omega)} \tag{1.12}$$

并称为**几何概率**。

几何概率具有如下性质：

(1) 非负性：对任一事件 A，$P(A)\geqslant 0$；

(2) 规范性：$P(\Omega)=1$；

(3) 完全可加性：设 $A_1,A_2,\cdots$为可列无穷个两两互不相容的事件，则

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

由完全可加性可得有限可加性，完全可加性是古典概率所不具有的性质。

例 1.8 （约会问题）两人相约于 7 点到 8 点之间在某地会面，先到者等候另一人 10 min，过时就离去，试求这两人能会面的概率。

解 以 x,y 分别表示两人到达某地的时刻，则

$$0\leqslant x\leqslant 60$$
$$0\leqslant y\leqslant 60$$

则样本空间 Ω 是由点(x,y)组成的边长为 60 的正方形，可以假设两人到达时刻在(0,60)内是等可能的，则点(x,y)落在 Ω 内是等可能的，因此，该问题为几何概型，两人能会面的充要条件是$|x-y|\leqslant 10$，能会面的点(x,y)的全体是位于区域 Ω 内且在$-10\leqslant x-y\leqslant 10$ 内部的区域，其面积为 60^2-50^2，故所求概率为

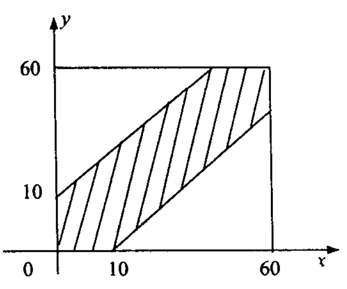


图 1-2 约会问题

$$p=\frac{60^2-50^2}{60^2}=0.3056$$

例 1.9 （蒲丰[Boffon]投针问题）在平面上画有等距离为 $a(a>0)$的一些平行线，向平面上任意投掷一长为 $l(l<a)$的针，试求针与平行线相交的概率。

解 因 $l<a$，所以针与平行线至多交于一点，以 M 表示针的中点，x 表示针投到平面上时 M 与最近一条平行线的距离，φ 表示针与最近一条平行线的夹角，则 $0\leqslant x\leqslant\frac{a}{2}$，$0\leqslant\varphi\leqslant\pi$。

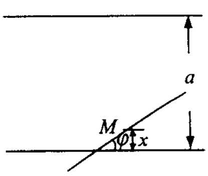


图 1-3 蒲丰投针问题

针与平行线相交的充要条件是：$x\leqslant\frac{l}{2}\sin\varphi$

所以样本空间 $\Omega=\{(\varphi,x):0\leqslant\varphi\leqslant\pi,0\leqslant x\leqslant\frac{a}{2}\}$

事件“针与平行线相交”所含样本点构成的区域为

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

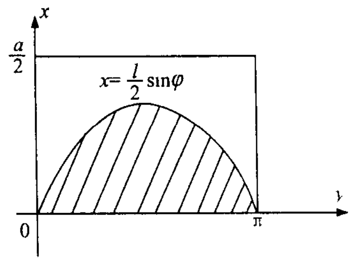


图 1-4　针与平行线相交的充要条件

所求概率 $p=\frac{G\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{\int_0^{\pi}\frac{l}{2}\sin\varphi d\varphi}{\frac{a}{2}\pi}=\frac{2l}{\pi a}$

由于最后的答案与 π 有关，因此人们利用它来计算 π 的数值，其方法是投针 N 次，观察针与平行线相交的次数 n，再以频率值 $\frac{n}{N}$ 作为概率 p 之值代入上式，求得

$$\pi=\frac{2lN}{an}$$

以上通过建立一个概率模型，它与某些我们感兴趣的量(这里是常数 π)有关，再经过适当的随机试验求出统计概率，从而可求得这个量，这类方法称为蒙特卡洛(Monte-Carlo)方法。

四、概率的公理化结构

以上我们讨论了怎样针对不同的问题分别用古典概率、几何概率和统计概率计算概率的方法，这些方法有一定的适用范围及某种局限性。因此从上述三种概率的任何一种出发来建立概率的理论都是不可能的。概率的公理化定义是从这三种概率定义的共同属性出发，指明概率应具有的基本性质，既包括这三种特殊情况，又具有更广泛的一般性。

由于定义概率所涉及的样本空间、事件的概念与集合的概念是一致的，因此我们先介绍集合论中有关 σ-域的概念，进而给事件及事件域以明确的数学定义。

1. 事件域

以集为元素的集合称为**集类**或**集合族**。

设 Ω 为一集合，$\mathscr{S}$ 是由 Ω 的某些子集组成的非空集类。

定义 1.5　设 $\mathscr{F}$ 是 Ω 上的一个非空集类，如果它满足下列条件：

(1) $\Omega\subset\mathscr{S}$；

(2) 若 $A\in\mathscr{S}$，则 $\overline{A}\in\mathscr{S}$；

(3) 若 $A_n\in\mathscr{F}$，$n=1,2,\cdots$，则 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{F}$，

则称 $\mathscr{F}$ 是 σ-域，或 σ-代数。

由定义 1.5 知，若 $\mathscr{F}$ 为 σ-域，则由(1)及(2)可得 $\varnothing\in\mathscr{F}$。

此外，若 $A_n\in\mathscr{F}$，$n=1,2,\cdots$，则

$$\bigcap_{n=1}^{\infty}A_n=\overline{\bigcup_{n=1}^{\infty}\overline{A}_n}\in\mathscr{F}$$

$$\bigcup_{n=1}^{k}A_n=A_1\cup A_2\cup\cdots\cup A_k\cup\varnothing\cdots\in\mathscr{F}$$

$$\bigcap_{n=1}^{k} A_n = A_1 \cap A_2 \cap \cdots \cap A_k \cap \Omega \cdots \in \mathscr{F}$$

可见，σ-域是关于可列次"和"、"积"及"差"运算封闭的集类。

定义 1.6 若 $\mathscr{F}$ 是由样本空间 Ω 的一些子集构成的一个 σ-域，则称 $\mathscr{F}$ 为**事件域**，$\mathscr{F}$ 中的元素(集合)称为**事件**。

可以验证，对于任意样本空间 Ω，它的所有子集(包括空集 $\varnothing$ 及 Ω)构成的集合族是一个 σ-域。当样本空间 Ω 中的样本点有限时，它的所有子集构成的 σ-域是一个有限集类。

今后，我们把任一样本空间 Ω，以及由 Ω 的子集所组成的一个 σ-域 $\mathscr{F}$ 写在一起，记为 $(\Omega,\mathscr{F})$，称为**可测空间**，$\mathscr{F}$ 中的每一个元素称为**可测集**。特别对有限样本空间，则称为**有限可测空间**。

2. 概率的公理化定义　概率空间

在很长时间内概率还不是一门成熟的数学学科，它的一些基本概念如概率的概念还没有明确的数学定义，到 19 世纪末人们开始注意概率论的公理化，曾提出过多种公理化系统。现在得到普遍承认的是柯尔莫戈洛夫(Колмогоров)公理化系统，它是柯尔莫戈洛夫在《概率论的基本概念》(1933 年发表)中提出来的。这个公理化系统综合了前人的成果，明确了概率的定义和概率论的基本概念，使概率论成为一门严谨的数学分支，对近几十年概率论的发展起了积极作用。

概率的公理化定义包括三个要素：样本空间、事件域和概率。在这个公理化体系中，概率是针对事件定义的，对应于事件域 $\mathscr{F}$ 中的每个元素 A，有一个实数 $P(A)$ 与之对应，这种从集合到实数的映射(记为 P)称为集合函数。因此，概率是定义在事件域上的集合函数。此外，还规定概率应满足的性质而不具体给出它的计算公式或计算方法。

下面给出概率的公理化定义。

定义 1.7 设 $(\Omega,\mathscr{F})$ 为可测空间，P 为定义在事件域 $\mathscr{F}$ 上的一个集合函数，如果它满足下列三个条件：

(1) 非负性：对于任意事件 $A\in\mathscr{F}$，$P(A)\geqslant 0$；

(2) 规范性：$P(\Omega)=1$；

(3) 完全可加性：若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$，且两两互不相容，则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \tag{1.13}$$

则称 P 为 $(\Omega,\mathscr{F})$ 上的概率。

定义 1.8 称三元总体 $(\Omega,\mathscr{F},P)$ 为概率空间，其中 Ω 为样本空间，$\mathscr{F}$ 是事件域，P 为概率。

对于每一个随机试验 E，都可以建立一个概率空间 $(\Omega,\mathscr{F},P)$ 与之对应，但究竟如何选定 Ω，怎样构造 $\mathscr{F}$，怎样规定 P，要根据具体情况而定，现举例说明。

例 1.10 (有限样本空间)设随机试验 E 的所有可能结果为有限个，这时 $\Omega=\{\omega_1,\omega_2,\cdots\omega_n\}$，事件域 $\mathscr{F}$ 可取为 Ω 的全体子集，$\mathscr{F}=\{\varnothing,\{\omega_1\},\cdots,\{\omega_n\},\{\omega_1,\omega_2\},\cdots,\{\omega_{n-1},\omega_n\},\cdots,\{\omega_1,\cdots,\omega_{n-1}\},\cdots,\Omega\}$，共有 2^n 个元素，$\mathscr{F}$ 为 σ-域。下面建立概率 P：

取一组实数 $\{a_i\}$，$a_i\geqslant 0$，$i=1,2,\cdots,n$，$\sum_{i=1}^{n} a_i=1$，令

$P(\omega_i)=a_i$，$i=1,2,\cdots,n$

$$P(A)=\sum_{\omega\in A}P(\omega_i)=\sum_{\omega\in A}a_i\ ,A\in\mathscr{F}$$

易证，P 为概率。所以$(\Omega,\mathscr{F},P)$为相应于随机试验 E 的概率空间。

例 1.11 （离散样本空间）设随机试验 E 的可能结果为可列个，这时样本空间 $\Omega=\{\omega_1,\omega_2,\cdots\}$；事件域 $\mathscr{F}$ 可取 Ω 中子集全体，$\mathscr{F}$ 为 σ-域；

取一列实数$\{a_i\}:a_i\geqslant0,i=1,2,\cdots,\ \sum_{i=1}^{\infty}a_i=1$，令

$$P(\omega_i)=a_i,i=1,2,\cdots,P(A)=\sum_{\omega_i\in A}P(\omega_i)=\sum_{\omega\in A}a_i\ ,A\in\mathscr{F}$$

则 P 为概率，$(\Omega,\mathscr{F},P)$为概率空间。

显然，离散概率空间包含有限概率空间。

五、概率的性质

假设已给一概率空间$(\Omega,\mathscr{F},P)$。从概率定义中的三条公理可以推出概率的下列性质。

1. $P(\varnothing)=0$

证明 因为$\varnothing=\varnothing\cup\varnothing\cup\cdots$，由概率的完全可加性得

$$P(\varnothing)=P(\varnothing)+P(\varnothing)+\cdots$$

因而得

$$P(\varnothing)=0$$

2. 有限可加性：对于任意 $A_i\in\mathscr{F},i=1,2,\cdots,n$，若 $A_iA_j=\varnothing,i\neq j,i,j=1,2,\cdots n$，则

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)$$

证明 令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$，由性质 1 $P(\varnothing)=0$，再由概率的完全可加性得

$$P(\bigcup_{i=1}^{n}A_i)=P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i)=\sum_{i=1}^{n}P(A_i)$$

这一性质说明，由概率的完全可加性可以推得概率的有限可加性，但逆命题不一定成立。

3. $P(\overline{A})=1-P(A)$；

证明 因 $A\cup\overline{A}=\Omega$ 且 $A\cap\overline{A}=\varnothing$，由性质 2 得 $P(A)+P(\overline{A})=P(\Omega)=1$

则

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

4. 单调性：如 $B\subset A$，则 $P(A-B)=P(A)-P(B)$，从而有 $P(B)\leqslant P(A)$，一般地 $P(A-B)=P(A)-P(AB)$

证明 若 $B\subset A$，则 $A=B\cup(A-B)$且 $B\cap(A-B)=\varnothing$，则有

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

得

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

又因为 $P(A-B)\geqslant0$，则 $P(B)\leqslant P(A)$

一般地，$A=A\cap(B\cup\overline{B})=AB\cup A\overline{B}$（见图 1-5），由 $B\cap\overline{B}=\varnothing$，得 $AB\cap A\overline{B}=\varnothing$，则有

$$P(A)=P(AB)+P(A\overline{B})=P(AB)+P(A-B),$$

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

5. 加法公式：$P(A\cup B)=P(A)+P(B)-P(AB)$；

次可加性：$P(A\cup B)\leqslant P(A)+P(B)$；

一般地有：$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)$

$+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$

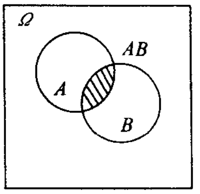


图 1-5　$A\cup B$

及　$P(\bigcup_{i=1}^{n}A_i)\leqslant\sum_{i=1}^{n}P(A_i)$　(1.14)

证明　$A\cup B=A\cup(B-AB)$（见图 1-5），且 A 与 $(B-AB)$ 互不相容，得

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

$$P(A\cup B)\leqslant P(A)+P(B)$$

对于一般情形的(1.14)式，可用数学归纳法证明，此处从略。

6. 概率具有连续性

(1) 先介绍单调事件列及其极限

设有一事件列 $A_1,A_2,\cdots$，若 $A_1\subset A_2\subset\cdots$，则称该事件列为递增的；若 $A_1\supset A_2\supset\cdots$，则称它为递减的。递增事件列和递减事件列统称为单调事件列。

若$\{A_n\}$递增，则$\lim\limits_{n\to\infty}A_n=\bigcup_{n=1}^{\infty}A_n$；

若$\{A_n\}$递减，则$\lim\limits_{n\to\infty}A_n=\bigcap_{n=1}^{\infty}A_n$；

称$\lim\limits_{n\to\infty}A_n$ 为单调事件列$\{A_n\}$的极限。

例如，考虑随机试验 E：向直线上随机投掷一质点，观察事件 $A=\{$质点落入$[0,1]\}$，$B=\{$质点落入$(0,1)\}$，$A_n=\{$质点落入$(0-\frac{1}{n},1+\frac{1}{n})\}$，$B_n=\{$质点落入$[0+\frac{1}{n},1-\frac{1}{n}]\}$$(n=3,4,\cdots)$。那么，易见$\{A_n\}$是递减事件列，$\lim\limits_{n\to\infty}A_n=\bigcap_{n=1}^{\infty}A_n=A$，$\{B_n\}$是递增事件列，$\lim\limits_{n\to\infty}B_n=\bigcup_{n=1}^{\infty}B_n=B$。

(2) 连续性定理

定理 1.1　设$\{A_n\}$，$A_n\in\mathscr{F}$，$n=1,2,\cdots$，是任一单调事件列，$\lim\limits_{n\to\infty}A_n=A$，那么

$$\lim_{n\to\infty}P(A_n)=P(\lim_{n\to\infty}A_n)=P(A)$$

概率的这一性质称为概率的连续性。

证明 1) 设$\{A_n\}$是一单调递增的事件列，易见

$$\lim_{n\to\infty}A_n=\bigcup_{n=1}^{\infty}A_n=\bigcup_{n=1}^{\infty}B_n$$

其中 $B_n=A_n-A_{n-1}$，(令 $A_0=\varnothing$)，$n=1,2,\cdots$，显然$\{B_n\}$两两互不相容。由概率的完全可加性，有

$$P(A)=P(\lim_{n\to\infty}A_n)=P(\bigcup_{n=1}^{\infty}A_n)=P(\bigcup_{n=1}^{\infty}B_n)$$

$$=\sum_{n=1}^{\infty}P(B_n)=\sum_{n=1}^{\infty}[P(A_n)-P(A_{n-1})]=$$