高等学校试用教材

概率论与数理统计简明教程

李贤平　卞国瑞　吴立鹏

高等教育出版社

高等学校试用教材

概率论与数理统计简明教程

李贤平　卞国瑞　吴立鹏

高等教育出版社

内 容 提 要

本书包括了概率论基础、马尔可夫过程和数理统计等三部分的基本内容，并结合计算机专业的特点在选材上进行了适当的删减与更新，该书内容符合国家教委制订的教学大纲的要求，可满足计算机等专业的教学需要。全书概念清楚，叙述简练，可作为高等学校计算机专业及相近专业的教材，也可供有关科技人员阅读参考。

本书前三章由卞国瑞、吴立鹏编写，后四章由李贤平编写，曾在复旦大学各专业试用达六年之久，最后由李贤平定稿。

高等学校试用教材

概率论与数理统计简明教程

李贤平　卞国瑞　吴立鹏

*

高等教育出版社 出版

新华书店上海发行所发行

商务印书馆上海印刷厂印装

*

开本 850×1168 1/32　印张 10.25　字数 244,000

1988 年 10 月第 1 版　1989 年 2 月第 1 次印刷

印数 0001—3,500

ISBN 7-04-001028-3/O·662

定价 2.40 元

目　　录

第一章　事件与概率

§1　随机现象与随机事件

一、随机现象

概率论是研究随机现象数量规律的数学分支.

人们在研究自然界和人类社会中各种事物运动的变化规律时会发现两类很不相同的现象.一类现象是在一定的条件之下,事物运动变化的规律是确定的. 一旦认识了这些规律就可以事先作出正确的预言. 例如太阳从东方升起;带电的物体之间,总是同性相斥异性相吸;在理想条件下,自由落体总是遵循着 $s=\frac{1}{2}gt^2$ 的运动规律;生物总是要经历生长、发育、衰老直至死亡的各个阶段等等. 这一类现象我们称之为**决定性现象**,把在一定条件下必然会发生的事情称之为**必然事件**. 反之,在一定条件下必定不会发生的事情称为**不可能事件**.

但是,在自然界和人类社会中还广泛地存在着与决定性现象有着本质不同的另一类现象. 例如抛掷一枚硬币事先并不能正确地预言结果是出现正面或反面;打桥牌时事先无法预料是否能分到有四张 A 的一副牌;事先也无法确定未来一小时内电话交换台接到的电话呼叫次数;同一条生产线上用同样的工艺生产出来的灯泡寿命长短也呈现出偶然性,这是由于在生产过程乃至对灯泡寿命试验的过程中种种偶然性的条件差异,使得每只灯泡的寿命事先不可能确定;在用仪器测量两点之间的距离时,虽然距离是确定的量,但是由于测量仪器受到周围环境(例如温度、湿度等等)的

影响以及观察者生理和心理上的偶然变化，使得观察结果中包含有事先无法确定的测量误差；再如某计算机下一次故障发生的时刻，某商场次日的营业总额等等，也是人们无法事先确定的．上述所举各例的一个共同特点是：在基本条件不变的情况下，一系列试验或观察会得到各种不同的结果．换句话说，仅仅就一次试验或观察而言，它会时而出现这种结果，时而出现那种结果，呈现出一种偶然性，也就是说人们无法事先预言．这种现象称为随机现象．

在实际中有许多问题与随机现象有关，而这些问题的解决有待于对随机现象的规律性进行深入的研究，不妨举一例．

设想某个体报贩可按其需要向报纸发行部门得到批发．他“需要的数量”当然根据卖出的数量来决定．最理想的情形是：事先能知道当天可以售出的报纸数量，然后按此数量批发．但这是做不到的．因为当天将能出售多少份报纸是机遇性的，也就是说，它是一个随机现象，事先不可能确切知道能卖出多少份报纸．这就提出了这样的问题：这个报贩要确定一个怎样的批发数量才能得到最大的利润。不妨假定每卖出一份报纸可得 0.015 元的纯利，而每剩余一份报纸将亏损 0.01 元．可以想见，如果盲目地批发太多，他将会有亏本的危险，而如果过于谨慎，始终使报纸的批发数处于供不应求的情况，则他只能赚到“保守”的数目，失去了一部分可以得到的利润．为了获得最大利润，他需要确定一个批发数量。这个数量的确定需要我们去研究：(1) 报纸零售这个随机现象的数量规律；(2)对于最大利润应该赋予随机意义下的明确概念。对于这种问题，我们已经学过的数学知识是不够的。而概率论为解决这种问题提供了方法。

尤其值得重视的是在商业等许多部门中都存在着类似上述的问题．由此可见，以随机现象为研究对象的概率论学科是有它的生命力的。

由于随机现象在一次观察中完全呈现出一种偶然性，因而要研究随机现象的规律，必须对随机现象进行大量重复的试验或观察，这就要求随机现象(或试验)本身可以在相同条件下重复地进行．当然，对于“相同条件”在数学上是把它理想化了的．在实际运用上只要求条件基本上相同即可．因为事实上绝对的“相同条件”是不可能存在的．

随机现象中可能出现的结果称为随机事件，简称为事件，例如上述报贩问题中，“日销售份数超过 100 份”就是一个随机事件，“日销售份数为 3000 份”也是一个随机事件．以后我们用 A, B, C, … 等大写拉丁字母来表示随机事件．

一个随机事件在一次试验或观察中可能发生也可能不发生，虽然呈现出一种偶然性，但在大量重复试验或观察中却呈现出明显的规律性，这就是频率的稳定性．

如果对于随机事件 A，在 N 次试验中发生了 n 次，则称

$$F_N(A)=\frac{n}{N}$$

为随机事件 A 发生的频率．

所谓频率稳定性是指，当试验次数 N 很大时，随机事件 A 发生的频率总是在某个固定常数的附近摆动．这一结论已为实践和理论两方面所证实，下面我们举一些体现频率稳定性的著名例子．

［例 1］ 抛一枚硬币，可能出现正面，也可能出现反面，事先要作出确定的判断是不可能的．但是假如硬币是均匀的，那我们当然有理由认为出现正面和反面的可能性应该一样大，即大量试验中出现正面和出现反面的频率都应接近于 50%．为验证这一点，历史上曾有不少人作过试验，其结果如下：

实验者	掷硬币次数	出现正面次数	频率
蒲丰(Buffon)	4040	2048	0.5069
皮尔逊(Pearson)	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

[例2] 任意指定一本英文书,翻到任意一页,任意指定该页中某一行及该行中任意一个位置,记录所得到的结果. 大量重复进行这一试验,可以发现26个字母及空格(空格指书中的空格和各种标点符号)被使用的频率相当稳定,下表是人们经过大量试验后得出的:

字母	空格	E	T	O	A	N	I	R	S
频率	0.2	0.105	0.072	0.0654	0.063	0.059	0.055	0.054	0.052
字母	H	D	L	C	F	U	M	P	Y
频率	0.047	0.035	0.029	0.023	0.0225	0.0225	0.021	0.0175	0.012
字母	W	G	B	V	K	X	J	Q	Z
频率	0.012	0.011	0.0105	0.008	0.003	0.002	0.001	0.001	0.001

研究字母使用频率,对于打字机键盘的设计(在便于操作的地方安排使用频率较高的字母键)、印刷铅字的铸造(使用频率高的字母相应多铸一些)、信息的编码(常用字母用较短的码)以及密码的破译等方面都是十分有用的.

为了说明这一点,我们选择英国生物统计学家Francis Galton的一段语录,原文是

Some people hate the very name of statistics, but I find them full of beauty and interest. Whenever they are not brutalized, but delicately handled by the higher methods, and are warily interpreted, their power of dealing with complicated phenomena is extraordinary. They are the only tools by which on opening can be cut through the formidable thicket of difficulties that bars the path of those who pursue the science of man.

这段语录的字母和空格的总数共计421个,各字母出现的频

数和频率如下表:

字母	空格	E	T	H	A	I	O	N	R
频数	72	49	38	27	25	24	22	20	20
频率	0.1710	0.1164	0.0903	0.0641	0.0594	0.0570	0.0523	0.04750	0.0475
字母	D	L	S	C	F	U	Y	P	B
频数	14	14	13	11	11	10	10	9	9
频率	0.0333	0.0333	0.0309	0.0261	0.0261	0.0238	0.0238	0.0213	0.0213
字母	M	W	G	V	K	X	Z	J	Q
频数	8	6	4	2	1	1	1	0	0
频率	0.0190	0.0143	0.0095	0.0038	0.0024	0.0024	0.0024	0	0

注意到这是一段不很长的文字，各个字母出现的频率与前表中的频率可能会有较大的差异,但统计结果表明,无论从频率的大小或顺序来看,都相当一致.

假定有人将某篇英文中的每个字母按一定规则进行变换，对于不知其规则的人，这段文字将成为不可辨认的“密码”. 但是按照字母出现频率的稳定性，人们还是可以将其中某些符号所代表的字母辨认出来的,加上运用英语的文字结构和语法规则,破译密码也就成为可能的了.

这两个例子表明，虽然随机现象在一次试验或一次观察中出现什么结果是偶然的，但在大量重复的观察和试验中却表明随机现象仍有着自己的规律性，即频率的稳定性. 这种规律性是由被观察对象的固有属性所决定的. 抛掷一枚硬币可以出现正面或反面,但硬币是均匀的,这一点决定着当多次抛掷钱币时,出现正面的次数应十分接近于总抛掷次数的 1/2; 至于英文字母被使用的频率,虽然各个作家的写作风格各不相同,书籍出版也各有自己的排版方式，但字母被使用的情况仍然受着文字结构、语法结构等

等固有规则的支配，英语的这种固有规则就决定了英文字母使用频率的稳定性.

随机事件的这种频率稳定性，我们称之为统计规律性. 只要我们对随机现象做较为深入的观察，我们就会发现统计规律性是十分普遍的现象.

下雨时，雨点是随机地落向地面的，在一个适当的地区之内，可以用某一地点的降雨量来表示该地区的降雨量，这就是频率稳定性的一个运用.

法国著名数学家拉普拉斯(1749—1827)，在他的时代里曾对男女婴儿的出生率进行过深入的研究，他发现男婴的出生率始终在22/43这个数值上摆动.

此外，在相同条件下耕种的庄稼，各块土地上的单位面积产量总是在平均单位面积产量附近摆动，而且具有某种对称性，同时还可以发现，在平均单位面积产量附近的集中一些，偏离大的就少一些. 一个射手向着目标射击，随着射击次数的不断增加，弹落点的分布就呈现出一种规律性：各个弹落点关于射击目标大致对称，偏离目标中心远的弹落点比偏离目标中心近的弹落点要少等等. 又如在一定条件下，一定时间内来到电话交换台的电话呼叫次数也呈现某种统计规律性.

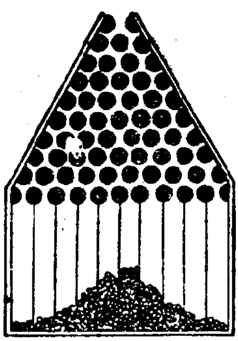

图1 高尔顿板

[例3] 另一个验证频率稳定性的著名试验是由上述语录的作者高尔顿(Galton)设计的. 它的试验模型如图1所示，自上端放入一小球任其自由落下，当小球碰到钉子时从钉子的左边或右边落下的机会相等，碰到下一排钉子时也是如此，最后落入底板中的某一格子. 因

此一个小球从上端放入最后落到底板中哪一个格子，事先是无法确定的．但实验证明，如大量放入小球，最后从底板上堆积起来的小球所形成的曲线形状，几乎总是一样的，也就是说小球落入底板中某一格子内的频率是十分稳定的．这个对于频率稳定性十分有说服力的试验模型，称为高尔顿板．

综上所述，随机事件的频率稳定性表明一个随机事件发生的可能性大小，是随机事件本身所固有的属性，它是可以度量的．也就是说，对一个随机事件 A，存在着一个与 A 相应的数 $P(A)$ 来刻划随机事件 A 发生的可能性大小．我们称 $P(A)$ 为事件 A 发生的概率．

对于随机现象，仅讨论它可能出现什么结果并无太大价值，而必须在讨论可能出现各种结果的同时，指出出现各种结果的可能性的大小，才有意义．只有这样，才能对随机现象作定量研究．

从上述可见，概率与频率有着十分密切的联系，当试验次数 N 充分大时，事件 A 发生的频率与概率应有

$$F_N(A) \approx P(A)$$

而且可以想见，概率与频率的性质应该基本相同，所以有必要考虑一下频率的性质：

(i) 非负性：对任意事件 A 有

$$F_N(A) \geqslant 0 \tag{1}$$

(ii) 以 Ω 表示必然事件，则由于必然事件在每次试验中必然发生，因此

$$F_N(\Omega) = 1 \tag{2}$$

(iii) 可加性：若 A 与 B 是两个事件，它们在一次试验中不会同时发生，以 $A+B$ 表示"A, B 中至少有一个发生"这个事件，则容易证明

$$F_N(A+B) = F_N(A) + F_N(B) \tag{3}$$

最后，根据频率的稳定性，自然会想到，当试验次数 N 无限增加时，频率与概率应具有某种极限关系. 在历史上它一直是概率论研究的一个重大课题，以后将会看到，这个结论的确成立，只是尚须对此问题的提法进一步明确化.

二、样本空间

下面我们将逐步引入概率论的一些基本概念.

为了叙述方便，把对随机现象进行的试验或观察统称为**随机试验**.

要认识一个随机试验，首先需要弄清楚它可能出现的各种结果. 例如，掷一枚硬币，可能出现的结果为正面或反面，连续掷两次(把它看成一次试验)可能出现的结果为(正，正)、(正，反)、(反、正)、(反，反)，这是这一试验所可能出现的全部基本结果，且在一次试验中它们不会同时出现. 这种可能出现的基本结果称为**样本点**，一般用 ω 来表示. 样本点全体构成的集合称为**样本空间**，用 Ω 表示. 对于一个具体的问题，确定一个相应的样本空间是研究随机现象的第一步.

[**例 4**] 一正立方体，六个面分别涂以红、黄、蓝、白、黑、绿六种颜色，任意抛掷一次，观察其朝上一面的颜色，则可能出现的全部结果为 $\Omega=\{$红、黄、蓝、白、黑、绿$\}$，此处，样本点为六种颜色中的一种. 如果把这正方体做成骰子，则 $\Omega=\{1, 2, 3, 4, 5, 6\}$，这两个样本空间 Ω 在本质上应该是一样的. 因此，对于有六个样本点的样本空间，可统一抽象地记为 $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$.

一般地，对于只有有限个样本点的随机试验，其样本空间可表示为 $\Omega=\{\omega_1, \omega_2, \cdots, \omega_n\}$.

样本空间的这种抽象表示，实际上是抓住了随机现象的本质，从而使研究的结果能适合于那些表面上不同而本质上一致的各种随机现象. 例如，只包含有两个样本点的样本空间，既能作为掷硬

币出现正面、反面的模型，也能用于描述产品检验中出现“合格”及“不合格”的模型，以及打靶时“命中”与“不中”的模型等等. 由于这样的原因，今后我们要用抽象化的随机模型作为同一类随机试验的代表.

只包含有限个样本点的样本空间称为**有限样本空间**.

[例5] 考察某计算中心在未来某一段时间内，所收到的来自各终端的请求次数，其可能结果为某一非负整数，且次数可能很大，难以规定一个合适的上界. 因此，取样本空间 $\Omega=\{0, 1, 2, \cdots\}$ 较为适当，这个样本空间包含有无穷多个样本点，但它们可以按某种次序全部排列出来，所以我们称之为**可列样本空间**，其一般形式为 $\Omega=\{\omega_1, \omega_2, \cdots\}$.

可列样本空间和有限样本空间统称为**离散样本空间**.

[例6] 定点计算中的舍入误差，是解题精度分析中的一个重要问题. 对于不同的问题或不同的数据，由于舍入所造成的计算误差也是不同的，可以将这种误差看作是随机的，它的取值范围可以充满某一区间 $[-a, a]$（a 的大小与计算机的字长有关）. 这时，相应的样本空间 $\Omega=[-a, a]$，而每个样本点 ω 是 $[-a, a]$ 中的某个数. 我们知道区间 $[-a, a]$ 上的数是不能逐个排列出来的，因此，它不是离散样本空间.

[例7] 打靶时，可以把靶面看成为一个无限平面，每个弹落点即样本点 ω 可以用坐标 (x, y) 表示，因此样本空间可取为 $\Omega=\{(x, y): -\infty<x, y<+\infty\}$.

由此可见，由于所考察的随机试验不同，因而相应的样本空间可能很简单也可能很复杂. 但是即使在同一随机试验中，由于所关心的问题不同，对于样本空间也可以有不同的选取. 例如打靶时，如研究的问题是弹落点所得的环数，样本空间可取为 $\Omega=\{0, 1, 2, \cdots, 10\}$，这样就把问题大大简化了. 因此，对于具体问题，怎

样选取一个恰当的样本空间也是值得研究的.

三、随机事件

事件是概率论最基本的概念之一,有了样本空间,我们可以把事件的概念进一步明确化.

掷一颗骰子, $\Omega=\{1, 2, 3, 4, 5, 6\}$. 以 A 表示"掷一颗骰子出现偶数点",显然 $A=\{2, 4, 6\}$; B 表示"不超过 4 点",则 $B=\{1, 2, 3, 4\}$,这里 A, B 都是随机事件, 它们由 Ω 中的部分样本点所组成.

在计算的舍入误差一例中,样本空间 $\Omega=[-a, a]$,以 A 表示"舍入误差的绝对值不超过 10^{-8}",则 $A=[-10^{-8}, 10^{-8}]$.

由此可见,所谓**事件**就是由样本点组成的某个集合,或者说是样本空间的某个子集.

在样本空间 Ω 上确定了某个集合 A, 亦即随机事件 A, 则对于 Ω 中每一样本点 ω 来说, 或者 ω 属于 A 或者不属于 A, 两者必居其一,分别记作 $\omega\in A$ 及 $\omega\overline{\in}A$. 当随机试验出现的结果 $\omega\in A$ 时称为事件 A **发生**. 反之,说一个随机试验发生了事件 A,则意味着 A 所包含的某个样本点 ω 恰为随机试验的结果. 这样, 就把事件 A 与 Ω 上的相应的子集 A 完全等同起来了. 由此可知,样本空间 Ω 也可看作一个事件, 由于它在每次试验中必然发生, 所以 Ω 就是**必然事件**. 类似地把不包含任何样本点的集合即空集也作为一个事件, 称为**不可能事件**, 记为 ϕ. 把 Ω 和 ϕ 看成是随机事件的极端情形, 对于今后进一步的讨论是必需的同时也是方便的.

把事件的定义与样本点的某个集合的概念联系起来是很重要的, 这使得我们能够从集合论的角度来研究事件. 下面在叙述事件之间的相互关系时, 我们将同时指出相应的集合之间的相互关系, 值得提醒的是在理解它们之间的关系时不要忘记随机事件的本来含义.

在一个样本空间中，同时考察几个事件，分析事件之间的相互联系，这不仅有助于我们认识事件的本质，而且对于从一些简单事件的概率推算出复杂事件的概率，以简化一些复杂事件的概率计算也是必需的，为此，我们下面将讨论事件之间的相互关系和事件的运算.

(1) 特款. 设有事件 A 和事件 B，在一次随机试验中，如果发生了事件 A，则必定发生事件 B，也就是说 A 的发生必然导致 B 的发生，则称事件 A 是事件 B 的**特款**，或称事件 B 包含了事件 A. 记为 $A\subset B$ 或 $B\supset A$. 这时事件 A 所包含的样本点都属于 B.

例如，以 A 表示事件"一小时之内来到交换台的电话呼叫次数是 6 的倍数"，而 B 表示事件"一小时之内来到交换台的电话呼叫次数为偶数". 则事件 A 发生必定导致事件 B 发生，所以 A 是 B 的特款. 此时 A, B 所包含的样本点分别是 $A=\{6, 12, 18, \cdots\}$, $B=\{2, 4, 6, \cdots\}$，显然，A 的样本点都属于 B，即 $A\subset B$.

容易理解，对于任一事件 A，有 $\Omega\supset A\supset\phi$.

(2) 等价. 如果对事件 A 和 B，同时成立 $A\supset B$ 和 $B\supset A$，则称事件 A **等价**于事件 B，记作 $A=B$. 实际上 A 和 B 是表示同一个事件，它们所包含的样本点完全相同.

(3)逆事件. 对于事件 A，"事件 A 不发生"也是一个事件，它称为事件 A 的**逆事件**，或称为 A 的**对立事件**，记作 $\overline{A}$. 它是由 Ω 中所有不属于 A 的样本点组成，如以 A 表示事件"来到的电话呼叫次数为偶数"，则它的对立事件 $\overline{A}$ 是"来到的电话呼叫次数为奇数". 由定义不难看出，对立事件是相互的，A 也是 $\overline{A}$ 的对立事件，因此有 $\overline{\overline{A}}=A$.

显然，Ω 和 ϕ 互为对立事件.

(4) 交事件. 若 A, B 是两个事件，"事件 A, B 同时发生"也是一个事件，称它为 A 与 B 的**交事件**，记为 $A\cap B$ 或 AB. 事件

AB 由同时属于事件 A 和 B 的样本点所组成. 显然, $A\cap\Omega=A$, $A\cap\phi=\phi$. 如以 A 表示"电话呼叫次数为偶数", B 表示"电话呼叫次数是 3 的倍数", 则 $A\cap B$ 表示"电话呼叫次数为 6 的倍数".

若 $A\cap B=\phi$, 它表示在同一次随机试验中, 事件 A 与事件 B 不会同时发生. 此时称 A 与 B **互不相容**, 互不相容的事件没有公共的样本点.

显然 A 与 $\overline{A}$ 是互不相容的.

(5) 并事件. 若 A, B 是两个事件, 则称事件"A 与 B 中至少有一个发生"为 A 与 B 的**并事件**, 记为 $A\cup B$. 它是由至少属于 A, B 中之一的全体样本点所组成. 显然, $A\cup\Omega=\Omega$, $A\cup\phi=A$. 如以 A 表示"电话呼叫次数不超过 100", B 表示"电话呼叫次数在 50 次到 150 次之间". 则它们的并事件 $A\cup B$ 为"电话呼叫次数不超过 150".

对于任何事件 A 和 B, 有 $A\cup B\supset A$ 和 $A\cup B\supset B$. 若 $A\supset B$, 则 $A\cup B=A$.

如果 A 与 B 互不相容, 记 $A\cup B=A+B$, 称为 A 与 B 的**和**. 在一次随机试验中, 如果 $A+B$ 发生, 则 A, B 之中只发生其中之一. 需要注意的是只有 A, B 互不相容时才能使用记号 $A+B$.

显然 $A+\overline{A}=\Omega$.

(6) 差事件. 若 A, B 是两个事件, 则称事件"A 发生而 B 不发生"为 A 关于 B 的**差事件**, 记为 $A-B$. 它是由属于 A 但不属于 B 的全体样本点所组成. 显然 $A-B=A\overline{B}$.

在进行运算时, 先进行逆的运算, 再进行交的运算, 最后进行并或差的运算.

正由于事件与集合概念之间的内在联系, 有时能在平面上用直观且方便的图形来表示事件之间的关系. 这种表示法称为文(Venn)图.