高等学校教材

概率论及数理统计

上 册

（第二版）

中山大学数学系
梁之舜　邓集贤
杨维权　司徒荣　邓永录
编　著

高等教育出版社

本书是中山大学数学力学系编《概率论及数理统计》(1980年)的修订第二版，现改为编者个人署名. 第二版与第一版相比有不少小的修改. 总的结构与第一版相同，仍分上、下册出版.

本书可作为综合大学、师范院校数学专业、应用数学专业和计算数学专业的教材，也可作其它有关专业的参考书.

学习本书只要求具有初等微积分基础知识，因此本书具有适应面广、便于自学的特点.

高等学校教材

概率论及数理统计

上 册

第二版

中山大学数学系

梁之舜　邓集贤　杨维权　司徒荣　邓永录　编著

*

高等教育出版社出版

新华书店北京发行所发行

文字六〇三厂印装

*

开本 850×1168　1/32　印张 12.375　字数 290 000

1988年5月第1版　1988年5月第1次印刷

印数 0001—5 620

ISBN7-04-000985-4/O·549

定价 2.50元

目　录

第一章　随机事件和概率

§1.1　随机事件的直观意义及其运算

一、必然现象与随机现象

在自然界里，在生产实践和科学实验中，人们观察到的现象大体可归结为两种类型．一类是可事前预言的，即在准确地重复某些条件下，它的结果总是肯定的，或是根据它过去的状态，在相同条件下完全可以预言将来的发展．我们把这一类型现象称之为确定性现象或必然现象．例如重物在高处总是垂直落到地面；在一个大气压下，水在 100°C 时会沸腾；水稻的生长从播种到收割，总是经过发芽、育秧、长叶、吐穗、扬花、结实这几个阶段；在射击时（假设空气阻力等可以忽略）弹道完全由射击的初始条件（如炮弹的初速、发射角和弹道参数等）决定．早期的科学就是研究这一类现象的规律性，所应用的数学工具如数学分析、几何、代数、微分方程等是大家所熟悉的．但人们逐渐还发现另一类型的现象，它是事前不可预言的，即在相同条件下重复进行试验，每次结果未必相同；或是知道它过去的状况，在相同条件下，未来的发展事前却不能完全肯定．这一类型的现象我们称之为偶然性现象或随机现象．如抛掷一个质地均匀的对称的硬币，结果可能是正面向上，或背面向上；新生的婴儿可能是男或是女；在相同海况与气象条件下，某定点海面的浪高时起时伏；当空气阻力等不能忽略时，弹道不能根据初始条件完全确定，可能向不同方向作程度不同的偏移，事前不能肯定．类似的例子还可以举出许多来．

初时人们把这种现象称为"偶然现象"是指它是"不正常的"、"出乎意料的"或者是"原因不明的"，甚至对于迅雷、疾风、陨石、地震等认为是天降的灾难.

是不是这些偶然现象都没有什么规律性可寻呢？事实上并非如此. 人们通过长期的反复观察和实践，逐渐发现所谓不可预言，只是对一次或少数几次观察或实践而言，当在相同条件下进行大量观察时，偶然现象都呈现某种规律，因而也是可以预言的. 例如根据各个国家各个时期的人口统计资料，新生婴儿中男婴和女婴的比例大约总是 1:1. 据传，在我国古代很早的时候就已经知道了这样一个结果. 又如人的高度虽然各不相同，但通过大量的统计，如果在一定范围内把人的高度按所占的比例画出"直方图"，就可以连成一条和铜钟的纵剖面一样的曲线；定点海面在一段时间内的浪高，也可以画出类似的曲线，如图 1.1.1. 还有更简单的例子(大家可以立即检验的)，均匀的硬币抛掷多次，正面和背面出现的次数比例总是近似 1:1，而且大体上抛掷次数愈多，愈接近这个比值. 历史上，蒲丰掷过 4040 次，得到 2048 次正面；皮尔逊掷过 24000 次，得到 12012 次正面.

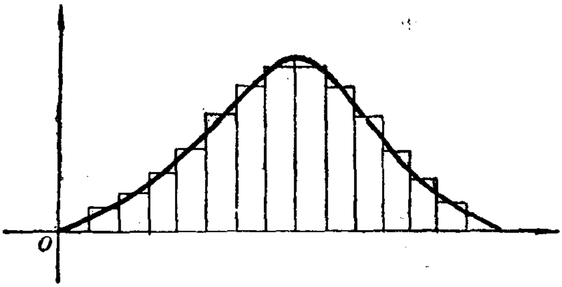

图 1.1.1

从上面的叙述中我们看到，自然界中存在着具有如下特性的现象：在一定的条件组实现时，有多种可能的结果发生，事前人们

不能预言将出现那种结果，但大量重复观察时，所得的结果却呈现某种规律，称为随机现象的统计规律性.

正如革命导师恩格斯所说："在表面上是偶然性在起作用的地方，这种偶然性始终是受内部的隐蔽着的规律支配的，而问题只是在于发现这些规律。"(《马克思恩格斯选集》中译本第四卷 243页).

概率论与数理统计就是研究随机现象统计规律性的一门数学学科.

根据马克思、恩格斯的论述，必然性与偶然性是对立统一的概念，偶然性不能理解为"碰巧的"，(为了避免作这种误解，人们往往把它称为"随机性")它蕴含内在必然性的规律；反过来，被断定为必然的东西，是由纯粹的偶然性构成的.

例 1.1.1 研究气体的性质知道，由于气体是由数目众多的分子构成，这些分子以很快的速度进行剧烈的运动，在运动的过程中相互碰撞而改变其动量和方向，因而每个分子的运动是随机现象．而大量的分子运动呈现出的总体现象——温度、压强却满足波义耳定律.

曾经有人认为，所以会出现事前不可预言的偶然现象，是因为我们对一个现象出现的原因还缺乏全面足够的知识，认为随着科学的发展和人类认识的深化，总有一天将不再存在不可预言的随机现象.

诚然，增加条件组的条件来减少随机性是可能的．例如，人的高度按不同的年龄、地区、性别是有差异的，若限定在同一地区，在一定年龄范围内选取同性别的人来测量高度，可能会得到比一般任意选取时有较均匀的结果，但随机因素的影响总是不可避免的．很多现象初始条件的稍微改变，其产生的后果却差别很大．在实际中即使人们有可能把条件组的条件绝对控制到每次都一样，但影响的因素还是大量的，且互相作用错综复杂．例如在例 1.1.1 关

于气体分子运动的例子中，即使我们把一立方厘米气体的 2.683×10^{18} 个分子的运动方程和初始条件都列出来，且不说我们要求出这组联立方程的解也是很困难的. 更应注意这里在建立方程和确定初始条件时已是忽略了许多次要的，但往往是随机的因素.

因此，偶然性现象是客观存在的，那种否认偶然现象的想法，是"力图用根本否认偶然性的办法来对付偶然性"(恩格斯《自然辩证法》中译本 196 页).

二、随机试验与事件

为了叙述方便，我们把对某种自然现象作一次观察或进行一次科学试验，统称为一个试验. 如果这个试验"在相同条件下可以重复进行"，而且每次试验的结果事前不可预言，我们就称它为一个随机试验.

下面，我们所说的试验都是指随机试验.

进行一个试验总有一个观察的目的，试验中会观察到有多种不同的可能结果. 例如抛掷一个质地均匀的硬币，我们的目的是要观察它那一面朝上，这里只有两种不同的结果："正面"或"背面"，至于硬币落在桌面上那一个位置，朝那个方向滚动等不在目的之列，我们不算作结果.

试验的每一个可能结果一般称为随机事件，简称为事件，我们用字母 $A,B,C,\cdots$ 表示.

例 1.1.2 在 $0,1,2,\cdots,9$ 十个数字中任意选取一个，可有十种不同的结果："取得一个数是 0"，…，"取得一个数是 9". 但还有其它可能结果："取得一个是奇数"，"取得一个是大于 4 的数"，"取得一个数是 3 的倍数"等等.

我们把不可能再分的事件称为基本事件. 例如在例 1.1.2中，"取得一个数是 0"，"取得一个数是 1"，…，"取得一个数是 9" 都是基本事件. 由若干基本事件组合而成的事件称为复合事件. 例如

“取得一个数为3的倍数”是一个复合事件，它由“取得一个数是3”，“取得一个数是6”，“取得一个数是9”三个基本事件组合而成.

一个事件是否称为基本事件是相对于试验的目的来说的. 例如量度人的身高，一般说，区间(0,4)中的任一实数，都可以是一个基本事件，这时，基本事件有无穷个；但如果量度高度是为了了解乘车是否需要买全票、半票或免票，这时就只有三个基本事件了.

三、事件的关系与运算

进行一个试验，有这样或那样的事件发生，它们各有不同的特性，彼此之间又有一定的联系. 下面我们引进事件之间的一些重要关系和运算，这将有利于今后对事件和它的概率的叙述和研究.

1. 在一定条件组下必然发生的事件称为必然事件. 在一定条件组下必然不发生的事件称为不可能事件. 必然事件用符号Ω表示，不可能事件用符号$\varnothing$表示. 把必然事件和不可能事件也算作随机事件，这对我们讨论问题是方便的.

例如，就目前世界上人高来说“人的高度小于4米”是必然事件，而“人的高度大于4米”则是不可能事件.

2. 如果两个事件A与B不可能同时发生，则称事件A与B为互不相容. 例如必然事件Ω与不可能事件$\varnothing$是互不相容的，又例如在例1.1.2中“取得一个数是0”和“取得一个数是1”是互不相容的.

如果$A_1, A_2, \cdots, A_n$中的任意两个事件是互不相容的，则称$A_1, A_2, \cdots, A_n$互不相容.

3. 如果事件A发生必然导致事件B发生，则称事件A含于事件B，或称事件B包含事件A，并记为$A \subset B$. 如例1.1.2中，令A表示“取得一数为4的倍数”，B表示“取得一数为偶然”，则$A \subset B$.

又如对任一事件 A，有 $\varnothing \subset A \subset \Omega$.

$A \subset B$ 的一个等价说法是，如果事件 B 不发生则事件 A 必然不发生.

如果 $A \subset B$ 且 $B \subset A$，则称 A 与 B 相等(或等价)，记为 $A = B$.

4. 如 C 表示“事件 A 与 B 中至少有一个发生”这一事件，称它为 A 与 B 的和(或并)，记为 $C = A \cup B$. 如在例 1.1.2 中，令 A 表示“取出一数为偶数”①，B 表示“取得一数大于 5”，则 $C = A \cup B$ 表示“取得一数或者大于 5，或者是偶数”，即等价于“取出一数为 0，2，4，6，7，8，9 中之一数”.

5. 如 D 表示“事件 A 与 B 同时发生”这一事件，称它为 A 与 B 的积(或交)，记为 $D = A \cap B$. 用上例，则 D 表“取得一数为 6 或 8”.

由事件积的定义，立即得到：

a) 对任一事件 A，有 $A \cap \Omega = A$，$A \cap \varnothing = \varnothing$；

b) 若 A_1，A_2 互不相容，则 $A_1 \cap A_2 = \varnothing$.

6. 如 E 表示“事件 A 发生而事件 B 不发生”这一事件，则称 E 为 A 与 B 之差，记为 $E = A - B$. 用前例，$E = A - B$ 表示“取得一个数为 0，2，4 中之一数”.

由二事件之差的定义立即得到：对任意事件 A 有

$$A - A = \varnothing; \quad A - \varnothing = A;$$
$$A - \Omega = \varnothing.$$

7. Ω 与 A 之差 $\Omega - A$ 这一事件称为 A 的逆事件，记为 $\bar{A}$，它表示“A 不发生”这一事件. 用前例，$\bar{A}$ 表示“取出一数为奇数”.

事件的和与事件的积都可以推广到有限(有穷)多个事件，即

$A = \bigcup\limits_{i=1}^{n} A_i$ 表示“A_1，A_2，…，A_n 中至少有一事件发生”这一事件.

① 为方便起见，以后均把数 0 算作偶数.

$B=\bigcap_{i=1}^{n} B_i$ 表示“$B_1, B_2, \cdots, B_n$ 同时发生”这一事件.

我们有时需要考虑可列无穷多个事件，需要把事件的和与积推广到可列无穷多个的情形，即考虑

$$A=\bigcup_{i=1}^{\infty} A_i \text{ 与 } B=\bigcap_{i=1}^{\infty} B_i,$$

记号 $\bigcup_{i=1}^{\infty} A_i$ 表示“可列无穷多个事件 A_i 中至少有一个发生”，而 $\bigcap_{i=1}^{\infty} B_i$ 表示“可列无穷多个事件 B_i 同时发生”.

为了说明这一需要，现举一例：一人进行科学实验，直到试验成功为止，若 A 表实验成功，A_i 表实验到第 i 次才成功，则显然有

$$A=\bigcup_{i=1}^{\infty} A_i.$$

可以验证一般事件的运算满足如下关系：

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

2) 结合律：$A\cup(B\cup C)=(A\cup B)\cup C$，
$A\cap(B\cap C)=(A\cap B)\cap C$；

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

分配律可以推广到有穷或可列无穷的情形，即

$A\cap(\bigcup_i A_i)=\bigcup_i(A\cap A_i)$，$A\cup(\bigcap_i A_i)=\bigcap_i(A\cup A_i)$；

4) $A-B=A\cap\bar{B}$；

5) 对有穷个或可列无穷个 A_i，恒有

$$\overline{\bigcup_i A_i}=\bigcap_i \bar{A}_i, \quad \overline{\bigcap_i A_i}=\bigcup_i \bar{A}_i.$$

四、用集合与几何图形表示事件，样本空间

对于事件及其运算，如果应用点集的概念和几何图示法，则较

直观而易于理解.

联系于每一随机试验的每一基本事件，用一个只包含一个元素ω的单点集$\{\omega\}$表示，由若干基本事件组成的复合事件，则用包含若干个元素的集合表示，由所有基本事件对应的全部元素组成的集合，给它一个特殊的名称，称为样本空间．由于任一随机试验的结果必然出现全部基本事件之一，这样，样本空间作为一个事件是必然事件．样本空间仍以Ω表示．我们称样本空间中的每一个元素为样本点．这样一来，集合论的知识（参阅附录II）就可以用来解释事件和事件的运算.

我们把它们的术语对照列表如下.

表 1.1.1

符　号	集　合　论	概　率　论
Ω	空　间	样本空间；必然事件
$\varnothing$	空　集	不可能事件
$\omega\in\Omega$	Ω中的点（或称元素）	样本点
$\{\omega\}$	单点集	基本事件
$A\subset\Omega$	Ω的子集A	事件A
$A\subset B$	集合A包含在集合B中	事件A含于事件B
$A=B$	集合A与B相等（或等价）	事件A与B相等（或等价）
$A\cup B$	集合A与B之和（或并）	事件A与B至少有一个发生（事件A与B之和或并）
$A\cap B$	集合A与B之交	事件A与B同时发生（事件A与B之积或交）
$\bar{A}$	集合A之余集	事件A的逆事件
$A-B$	集合A与B之差	事件A发生而事件B不发生（事件A与B之差）
$A\cap B=\varnothing$	集合A与B没有公共元素	事件A与B互不相容

如果以平面上的某一矩形表示样本空间，矩形内的每一点表示样本点，则事件的运算可通过平面上的几何图形表示．如果用两个小圆形表示事件A和B，如图1.1.2所示，阴影部分表示事件

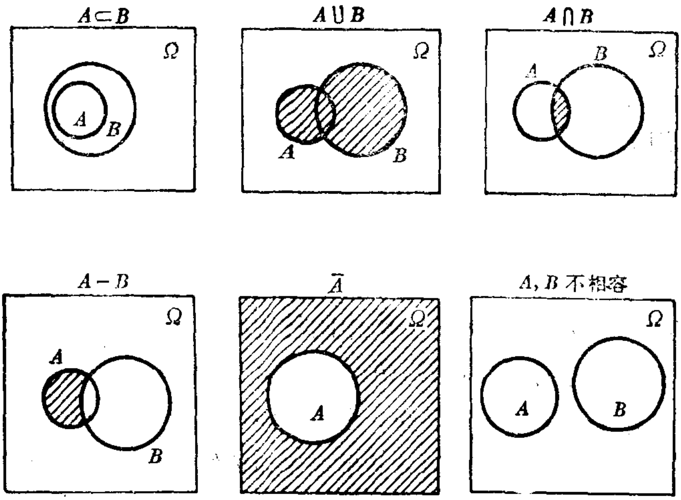

图 1.1.2 ($A\cup B$, $A\cap B$, $A-B$, $\bar{A}$ 分别为图中阴影部分)

A 与 B 的各种关系及运算.

上面说过样本空间是由所有代表基本事件的样本点构成，而基本事件则直接联系于每个随机试验，下面举一些例子说明:

如在例 1.1.2 中样本点是 $\omega_1=0, \omega_2=1, \cdots, \omega_{10}=9$.

基本事件为 $\{\omega_i\}$ $(i=1,2,\cdots,10)$;

样本空间 $\Omega=\{0,1,\cdots,9\}$.

例 1.1.3 设我们同时抛掷两枚硬币，其基本事件为

$$A_1=\{\omega_1\},\ A_2=\{\omega_2\},$$

$$A_3=\{\omega_3\},\ A_4=\{\omega_4\},$$

其中 $\omega_1=$(正面, 正面), $\omega_2=$(正面, 背面), $\omega_3=$(背面, 正面), $\omega_4=$(背面, 背面)[①]是样本点. 这时，样本空间由四个样本点构成，即

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}.$$

例 1.1.4 计算某电话站总机在 $(0,T]$ 内的呼叫次数，则基本

---

① (背面, 背面)表示第一个硬币出现背面，第二个硬币也出现背面，余类推.

事件为 $A_0=\{\omega_0\}, A_1=\{\omega_1\}, \cdots, A_n=\{\omega_n\}, \cdots$
其中 $\omega_k=(k$ 次呼叫$)(k=0,1,2,\cdots)$，样本空间为 $\Omega=\{\omega_k: k=0, 1,2,\cdots\}$.

例 1.1.5 设我们测量某一零件时，考虑其测量结果与真正长度的误差. 一般说，它可用一数 x 表示，基本事件为 $\{x\}(-\infty<x<\infty)$①，实数轴上每一点都是其样本点，整个数轴为其样本空间，即 $\Omega=\{x: x\in R_1\}=R_1$.

例 1.1.6 若在例 1.1.5 中同时考虑测量某零件的长度 x 和直径 y，基本事件为 $\{(x,y)\}(0<x<\infty,\ 0<y<\infty)$. 则每一样本点为坐标平面上第一象限的一个点，即数对 $(x,y)(0<x<\infty, 0<y<\infty)$. 这样，其样本空间就是坐标平面的第一象限，即 $\Omega=\{(x,y): 0<x<\infty, 0<y<\infty\}$.

§1.2 概率的直观意义及其计算

我们观察一个随机试验的各种事件，一般来说，总会发现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.

研究随机现象不仅要知道它可能出现那些事件，更重要的是要研究各种事件出现可能性大小，揭示出现这些事件的内在的统计规律，只有这样才有利于我们认识世界和改造世界. 例如知道了某电话总机在 24 小时内出现的呼叫次数的可能性大小，就可以根据要求配置一定的线路设施、管理人员等.

如果只能大概地比较事件发生的可能性大小，要想进行确切的推断，得出精确结论是不行的. 我们要求有一个刻划事件发生可能性大小的数量指标，这个数量指标至少应该满足两个要求：

(i) 它应具有一定的客观性，不能随意改变. 而且理论上可

① 全体实数所成的集记为 R_1.

以通过在"相同条件下"大量的重复试验予以识别和检验.

(ii) 它必须符合一般常情. 例如,事件发生可能性大的,它的值就大;事件发生可能性小的,它的值就小;必然事件的值最大,不可能事件的值最小而等于零.

我们把刻划事件发生可能性大小的数量指标叫做事件的概率. 事件 A 的概率以 $P(A)$ 表示,并且规定 $0 \leqslant P(A) \leqslant 1$.

在概率论的发展历史上,人们曾针对不同的问题,从不同的角度给出了定义概率和计算概率的各种方法. 然而所定义的概率都存在一定的缺陷,所以我们毋宁说它是一些计算概率的方法.

一、古典概率

以抛掷质地均匀的硬币为例,人们自然想到由于硬币两面是对称的,所以出现"正面"和"背面"的可能性都一样,人们有理由认为出现"正面"和"背面"的概率都是 $\frac{1}{2}$.

在古代较早的时候,在某些特殊情形下,人们利用研究对象的物理或几何性质所具有的对称性,确定计算概率的一种方法如下:对于某一随机试验,如果它的全体基本事件 $E_1, E_2, \cdots, E_n$ 是有穷的,且具有等可能性,则对任意事件 A,对应的概率 $P(A)$ 由下式计算:

$$P(A)=\frac{\text{事件 } A \text{ 包含的基本事件数}(k)}{\text{基本事件总数}(n)}, \qquad (1.2.1)$$

并把它称作古典概率.

例 1.2.1 设电话号码由 0, 1, 2, …, 9 共十个数字中任意五个数字组成,设某一户的电话号码是 51710, 问当不知道这个电话号码时,一次拨号就能拨对该电话号码的概率多少?

解 依题意全部电话号码有 10^5 个,当不知道电话号码时,拨 10^5 个电话中的任一电话号码的可能性是相等的. 令 A 表示"一次

拨号就能拨对该用户号码"这一事件，则这时全部基本事件数 $n=10^5$，而事件 A 只包含一个基本事件，即 $k=1$，按古典概率计算得 $P(A)=\frac{1}{10^5}=0.00001$.

可见当不知道该用户电话号码时，一次拨号就能拨对该电话号码的可能性是很小的.

例 1.2.2 100 只同批生产的外形一样、同型号的三极管中按电流放大系数分类，有 40 只属于甲类，60 只属于乙类. 在按下列两种抽取方法中，求下列事件的概率：

A="从 100 只中任意抽取 3 只，3 只都是乙类"；

B="从 100 只中任意抽取 3 只，其中有两只是甲类，1 只是乙类".

抽取方法是

(1) 每次抽取一只，测试后放回，然后再抽取下一只（这种抽取方法称为有返回抽样）；

(2) 每次抽取一只，测试后不放回，在剩下的三极管中再抽取下一只(这种抽取方法称为不返回抽样).

解 先求事件 A 的概率：

(1) 由于每次抽取测试后均放回，因此，每次都是从 100 个三极管中抽取. 从 100 只中任取 3 只的所有可能取法有 100^3 种，即基本事件总数

$$n=100^3.$$

现考虑事件 A 所包含的基本事件数. 由于属于乙类的三极管有 60 个，因此抽取 3 只都是乙类的所有可能取法有 60^3 种，即 A 包含的基本事件数为

$$k=60^3.$$

按古典概率计算得