高等学校教材

随机数据处理方法

编著　王清河　常兆光　曹晓敏　许晓婕

中国石油大学出版社

高等学校教材

随机数据处理方法

编著 王清河 常兆光 曹晓敏 许晓婕

本教材第二版于2002年获山东省省级教学成果三等奖

中国石油大学出版社

内 容 简 介

本书系统介绍处理随机数据的统计方法及统计方法的应用。内容包括随机事件与概率、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、数理统计初步、回归分析、随机过程、方差分析与正交试验设计、判别分析、聚类分析等。书中收录了许多实例。各章有例题及练习题，书末有附录。

本书可作为高等理工院校非数学类专业概率论与数理统计课程的教材，也可供应用统计的相关工作者和工程技术人员参考。

图书在版编目(CIP)数据

随机数据处理方法/王清河等编著. —4 版. —东营：中国石油大学出版社，2011. 2

ISBN 978-7-5636-3350-0

Ⅰ. ① 随…　Ⅱ. ① 王…　Ⅲ. ① 随机变量—数据处理　Ⅳ. ① 0211. 5

中国版本图书馆 CIP 数据核字(2010)第 004046 号

书　　名：随机数据处理方法

作　　者：王清河　常兆光　曹晓敏　许晓婕

责任编辑：魏　瑾

出 版 者：中国石油大学出版社（山东 东营　邮编 257061）

网　　址：http://www.upcbook.com.cn

电子信箱：weicbs@163.com

印 刷 者：沂南县汇丰印刷有限公司

发 行 者：中国石油大学出版社（电话 0546—8392062）

开　　本：185×260　印张：18.75　字数：480 千字

版　　次：2011 年 2 月第 4 版第 1 次印刷

定　　价：31.80 元

序

概率论与数理统计在各种领域中都有着广泛的应用，现已成为高校文、理、工、医、经济等学科学生的重要的必修课程。笔者过去也曾多年讲授这门课程，感到同学们在学习这门课程中的困难，不在于所用数学工具的高深，因为本课程在以讲授方法为主的范围内，所涉及的数学知识基本上限于微积分与线性代数的基础性部分。这门科学之所以难学难教，有两个原因：一是基本概念与学生们习惯的非随机数学相比，有较大差异，不容易领会，而在这一点上有欠缺，就会妨碍学生对方法的接受与理解。二是这门科学重在实用且涉及的题材方面广，方法多，许多方法的应用不能光靠套公式解决，它涉及模型和方法的选择、问题的恰当的提法及分析结果的正确解释等。这些都是在讲授和学习这门课程中不容易过而又必须过的"关"，需要教师和同学共同作出努力。为达到这一点，一本良好的教材有着根本的重要性。

坊间现在出版的这方面教材，包括国人自著和译作，为数不算少，也各有其优缺点和适用范围。最近见到常兆光教授等著《随机数据处理方法》(第三版)的原稿，感到是一部具有众多优点的教材，对克服笔者在前面指出的两个难点很有针对性。首先，此书用平实的语言对概率论与数理统计的基本概念作了细致而深入浅出的解释，并通过阐释性的例子、数字计算例子与实用例子，从多个角度加以仔细说明，这很有助于读者对基本概念的正确理解。其次，书中对各种常用的方法，通过应用实例，作了仔细而清楚的介绍，这些实例来自各个方面，显示了这门学科的广泛应用，这一点因书中所附的大量的带应用色彩的习题而得到加强。

本书的另一个突出优点是材料丰富，介绍了一些在基本课程中往往不多涉及，而在实用上很常用的一些领域和方法，这包括本书的第 8～10 章，这几章所介绍的内容有广泛的应用，但在关于本学科的基础性教材中，由于教学时间等方面的考虑，多未能包含。本书提供了这方面的材料，可以供教师在时间允许的情况下适当选用。另外，笔者一直赞成这样的观点：教科书不必与课堂讲授的内容完全一致，有些内容由于时间关系，未列入教学大纲，但因其重要性，学生又必要有所了解，这种内容可以写进教科书，它对学生的进一步学习，以及自学者，都是很有益的。

今年春天我在石油大学待了一个多月，当时常兆光教授等正在为第三版定稿而努力，在这一版中，作者根据教学实践中的经验，对内容作了不少的充实和改进，使之更适合于讲堂教学的需要及自学者的进修。相信本书的出版，定会受到广大读者欢迎，故特书此表示祝贺。

陈希孺

2004.12.17

前　言

概率统计是数学的一个分支，其方法已经广泛应用于多个领域，如地质、石油、气象、水文等。特别是近十几年来，由于电子计算机的普及，它的发展更加活跃和深入，其方法已普遍受到人们的重视。

概率统计方面的书已有很多，多数是将一元统计与多元统计分开。有的偏重于理论，有的偏重于方法，而以应用为主的很少介绍近代统计方法。我们采取的原则是：(1)基本概念和基本方法详细介绍，对工程中应用较广泛的方法着重从应用的角度介绍其应用背景、统计思想、应用条件及具体实现步骤；(2)强调工程应用，尽可能介绍现场应用统计方法的例子，提高读者用统计方法处理实际问题的能力；(3)在介绍基本统计方法的同时，尽可能介绍近年发展起来的一些新的统计方法；(4)对古典内容作适当压缩，使读者在有限的篇幅内能了解随机数据处理方法的概貌。

王才经教授审阅了本书的全稿，并提出了许多宝贵意见。本书的出版还得到了石油大学出版社的大力支持，我们在此一并表示衷心的感谢。

限于编者水平，书中不妥之处还请专家、同仁批评指正。

编　者

1992 年 9 月 22 日

再版前言

第三版出版后，我们曾收到许多专家、主讲教师及读者的热情指教和建议，结合教学工作的体会，借再版的机会对本书作了如下修改：

1. 根据课程要求，增加了随机过程的内容；

2. 为方便计算，在不影响理论和方法掌握的基础上，增加了 Excel 的内容，并对本书中涉及较复杂运算的例题用 Excel 函数进行了计算；

3. 分布函数的定义采用了右连续定义；

4. 对部分内容及表达方式进行了修改；

5. 对印刷错误作了校正；

6. 对多元统计部分内容作了较大改动，更能体现统计方法的应用性.

本书的再版修改过程得到了李元教授的大力支持和帮助，李元教授审阅了全书，并提出了若干修改意见. 常秦博士应用 Excel 及其他统计软件对书中大部分习题进行了较详细的解答，在此一并表示衷心的感谢.

竭诚地希望读者继续对本书提出批评建议.

编　者

2011 年 2 月

目　　录

第1章　随机事件与概率

在生产实践、科学实验和日常生活中，人们观察到的现象一般可分为两种类型.一类是确定性现象，即在一定条件下，一定发生或一定不发生的现象，也称为必然现象.例如：纯净水在一个标准大气压下，加热到100 ℃必然沸腾；上抛的物体必然下落等.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高等数学、几何、代数等是大家所熟悉的.但人们逐渐发现了另一类现象，它是事前不可预言的，即在一定条件下，可能发生也可能不发生的现象，这一类现象我们称之为偶然现象或随机现象.例如：抛掷一枚均匀硬币，是正面朝上还是反面朝上；从一批含有次品的产品中任取一件，取得的产品是正品还是次品；新生婴儿是男还是女等.这些都是事前不能肯定的.类似的例子还可以举出很多.

必然现象遵循必然性规律，人们根据已知的事实就可以推断它将发生的结果.随机现象具有明显的不确定性(随机性)，就一次试验(或观察)而言，其结果难以确定，但若进行大量重复试验，其结果就会呈现出某种规律性，即所谓统计规律.概率论与数理统计的任务就是要研究和揭示随机现象的这种统计规律性.

§1.1　随机试验与随机事件

1. 随机试验

在工农业生产、科学实验和现实生活中，我们遇到过各种各样的试验.在概率论中，我们把试验作为一种广泛的术语，它包括各种科学试验，甚至对某一事物的某一特征的观察也可认为是一种试验.例如(以 E 或 E 加一个下标表示试验)：

E_1：掷一枚均匀硬币，观察正面(记为 H)、反面(记为 T)出现的情况；

E_2：将一枚均匀硬币掷两次，观察正、反面出现的情况；

E_3 将一枚均匀硬币掷两次，观察正面出现的次数；

E_4：掷一颗均匀骰子，观察出现的点数；

E_5：一名射手进行射击，直到击中为止，观察射击次数；

E_6：从一批灯泡中任意抽取一只，测其寿命.

不难看出，以上几个试验具有下面三个共同特点：

(1) 试验可以在相同条件下重复进行；

(2) 每次试验的可能结果不止一个，但能事先明确试验的所有可能结果；

(3) 进行一次试验之前，不能确定哪一个结果会出现.

在概率论中，我们把具有以上三个特征的试验称为随机试验，简称为试验，记作 E.

2. 随机事件与样本空间

我们将试验 E 所有可能出现的结果组成的集合称为 E 的样本空间，记作 Ω. Ω 中的每个

元素(可能结果)称为样本点.例如,上述试验 $E_i(i=1,2,3,4,5,6)$的样本空间 Ω_i 分别为:

$\Omega_1=\{H,T\}$; $\Omega_2=\{HH,HT,TH,TT\}$; $\Omega_3=\{0,1,2\}$;

$\Omega_4=\{1,2,3,4,5,6\}$; $\Omega_5=\{1,2,3,4,\cdots\}$; $\Omega_6=\{t|t\geqslant 0\}$.

需要指出的是:样本空间中的样本点是由试验目的所确定的.例如,E_2 和 E_3 同是将一枚均匀硬币掷两次,但由于试验目的不同,其样本空间也就不一样.

样本空间包含了试验 E 的所有可能结果,我们将每一个可能的结果称为随机事件(简称为事件),通常用大写字母 A,B,C 等表示.只包含一个样本点的事件称为基本事件.例如,掷骰子试验中,每一个可能出现的点数都是基本事件.而有两个或两个以上的基本事件(样本点)组成的事件称为复合事件.例如,将一枚均匀硬币连续掷两次,观察正、反面出现的情况,其样本空间 $\Omega=\{HH,HT,TH,TT\}$.则两次都出现正面为一基本事件,记作 $A=\{HH\}$;至少有一次出现正面为一复合事件,记作 $B=\{HH,HT,TH\}$.可以看出 B 事件是由三个类似于 A 事件的基本事件所组成的,即这三个事件中只要有一个发生就认为 B 事件发生.

在每次试验中,一定发生的事件叫做必然事件,记作 Ω.而一定不发生的事件叫做不可能事件,记作$\varnothing$.

需要指出的是:无论是必然事件、随机事件还是不可能事件,都是相对"一定条件"而言的.条件发生变化,事件的性质也发生变化.例如,抛掷两颗骰子,"出现的点数之和为 5 点"及"出现的点数之和大于 5 点"都是随机事件.若同时抛掷 6 颗骰子,"出现的点数之和为 5 点"则是不可能事件,而"出现的点数之和大于 5 点"则是必然事件了.为了以后讨论问题方便,通常将必然事件和不可能事件看成是特殊的随机事件.

3. 事件及其运算关系

由样本空间的定义知,样本空间是随机试验中所有可能出现的结果(样本点)组成的集合.因此随机事件又可理解为样本空间中具有某种特性的样本点构成的集合,而基本事件可看成此集合的元素.这样一来,集合论中集合之间的运算均可推广到事件之间的运算中来.

若记 Ω 为样本空间(必然事件),$\varnothing$为不可能事件,e 为基本事件,$A,B,A_1,A_2,\cdots,A_n,\cdots$为随机事件.则有事件之间的运算关系如下:

(1) 包含关系:若 $A\subset B$,则称事件 B 包含事件 A,它表示事件 A 发生必导致事件 B 发生,如图 1-1 所示.

对任一事件 A,都有$\varnothing\subset A\subset\Omega$.

(2) 相等关系:若 $A\subset B$ 且 $B\subset A$,则称事件 A 等于事件 B,记作 $A=B$.

(3) 和事件:事件 $A\cup B$ 称为事件 A 与事件 B 的和事件,它表示事件 A 与事件 B 至少有一个发生,如图 1-2 所示.

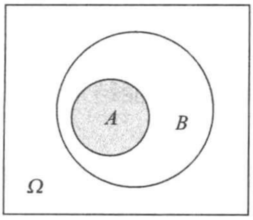

图 1-1

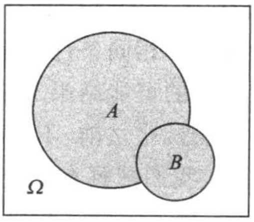

图 1-2

事件 $A_1\cup A_2\cup\cdots\cup A_n\triangleq\bigcup\limits_{i=1}^{n}A_i$ 称为有限个事件的和事件，它表示事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生.

类似地，事件 $A_1\cup A_2\cup\cdots\cup A_n\cup\cdots\triangleq\bigcup\limits_{i=1}^{\infty}A_i$ 称为无限多个事件的和事件，它表示事件 $A_1,A_2,\cdots,A_n,\cdots$ 至少有一个发生.

（4）积事件：事件 $A\cap B$ 称为事件 A 与事件 B 的积事件，简记为 AB，它表示事件 A 与事件 B 同时发生，如图 1-3 所示.

事件 $A_1\cap A_2\cap\cdots\cap A_n\triangleq\bigcap\limits_{i=1}^{n}A_i$ 称为有限个事件的积事件，它表示事件 $A_1,A_2,\cdots,A_n$ 同时发生.

类似地，事件 $A_1\cap A_2\cap\cdots\cap A_n\cap\cdots\triangleq\bigcap\limits_{i=1}^{\infty}A_i$ 称为无限多个事件的积事件，它表示事件 $A_1,A_2,\cdots,A_n,\cdots$ 同时发生.

（5）差事件：事件 $A-B$ 称为事件 A 与事件 B 的差事件，表示事件 A 发生而事件 B 不发生，如图 1-4 所示.

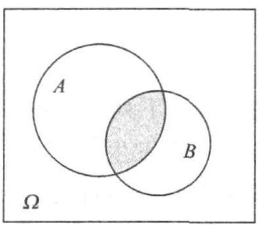

图 1-3

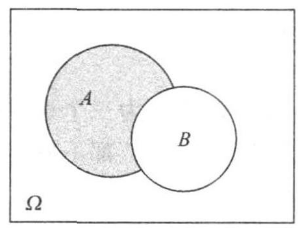

图 1-4

（6）互不相容（互斥）事件：若 $AB=\varnothing$，则称事件 A 与事件 B 互不相容，它表示事件 A 与事件 B 不同时发生，如图 1-5 所示.

基本事件是互不相容的.

类似地，若 $\forall\, i\neq j$，$A_iA_j=\varnothing\,(i,j=1,2,\cdots)$，则称 $A_1,A_2,\cdots,A_n,\cdots$ 为两两互不相容的.

若 $A_1,A_2,\cdots,A_n$ 两两互不相容，且在每次试验中事件 $A_1,A_2,\cdots,A_n$ 必发生其中之一，即 $\bigcup\limits_{i=1}^{n}A_i=\Omega$ 且 $A_iA_j=\varnothing\,(i\neq j)$，则称 $A_1,A_2,\cdots,A_n$ 为互不相容事件完备组.

（7）对立事件（逆事件）：若 $AB=\varnothing$ 且 $A\cup B=\Omega$，则称事件 A 与事件 B 互逆，又称事件 A 为事件 B 的对立事件（或事件 B 为事件 A 的对立事件），记作 $A=\overline{B}$（$\overline{B}$ 表示 B 不发生），$B=\overline{A}$，如图 1-6 所示.

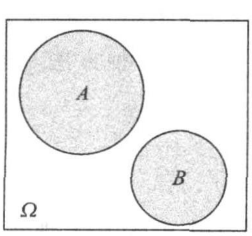

图 1-5

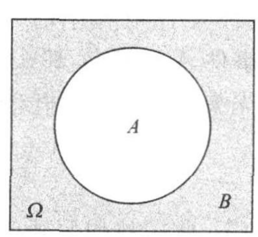

图 1-6

需要指出的是：由（6）、（7）知，若事件 A 与事件 B 互逆，则事件 A 与事件 B 一定互不相容，反之则不一定.

例 1-1 在试验 E_2 中，若记 $B_1=\{HH\}$，$B_2=\{HT\}$，$B_3=\{TH\}$，$B_4=\{TT\}$，则：

A_1 表示"第一次出现正面"，即 $A_1=\{HH,HT\}\triangleq B_1\cup B_2$；

A_2 表示"两次出现同一面"，即 $A_2=\{HH,TT\}\triangleq B_1\cup B_4$；

A_3 表示"只出现一次正面"，即 $A_3=\{HT,TH\}\triangleq B_2\cup B_3$.

那么，$A_1\cup A_2=\{HH,HT,TT\}$，$A_1\cap A_2=\{HH\}=B_1$，$A_1-A_2=A_1\overline{A_2}=B_2$. 由于 $A_2A_3=\varnothing$，故 A_2 与 A_3 互不相容，又由于 $A_2\cup A_3=\Omega$，所以 A_2 与 A_3 互逆.

(8) 德·摩根定律（对偶原理）：

$$\overline{A\cup B}=\overline{A}\cap\overline{B},\quad \overline{A\cap B}=\overline{A}\cup\overline{B};$$

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i},\quad \overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}.$$

(9) 运算规律：设 A,B,C 为三个事件，则有

交换律： $A\cup B=B\cup A,\quad A\cap B=B\cap A.$

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

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

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

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

例 1-2 设 A,B,C 为三个事件，试用事件之间的运算关系表示下列事件：

(1) A 发生而 B 与 C 都不发生； (2) A,B,C 恰有一个发生；

(3) A,B,C 至少有一个发生； (4) A,B,C 至多有两个事件发生.

解 (1) A 发生而 B 与 C 都不发生，也就是 $A,\overline{B},\overline{C}$ 同时发生，即 $A\cap\overline{B}\cap\overline{C}$ 或 $A-B-C$.

类似可得：

(2) $(A\cap\overline{B}\cap\overline{C})\cup(\overline{A}\cap B\cap\overline{C})\cup(\overline{A}\cap\overline{B}\cap C)$.

(3) $A\cup B\cup C$.

(4) A,B,C 至多有两个事件发生，相当于 $\overline{A},\overline{B},\overline{C}$ 至少有一个发生，即 $\overline{A}\cup\overline{B}\cup\overline{C}$；或 A,B,C 不同时发生，即 $\overline{ABC}$.

§1.2 频率与概率

一个随机试验有许多可能的结果，而在多数情况下，我们想知道的往往是随机事件（可能结果）发生的可能性大小. 如建造水坝，为确定坝高，需要知道建造水坝地段每年最大洪水到达某高度的可能性大小. 在概率论中，将描述随机事件 A 发生的可能性大小的数记为 $P(A)$，称 $P(A)$ 为随机事件 A 的概率. 那么如何确定随机事件 A 的概率呢？一种方法是通过反复试验来确定，为此先来讨论频率的概念.

1. 频率

[引例 1-1] 将一枚均匀硬币在同样条件下连续掷 n 次，用 A 表示出现正面这一事件，$n(A)$ 表示 n 次重复试验中 A 出现的次数，则 $\frac{n(A)}{n}$ 在一定程度上能反映事件 A 发生的可能性大小，将其记为 $f_n(A)\triangleq\frac{n(A)}{n}$. 试验结果见表 1-1 及表 1-2.

表 1-1

试验序号	$n=5$		$n=50$		$n=500$	
	$n(A)$	$f_n(A)$	$n(A)$	$f_n(A)$	$n(A)$	$f_n(A)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

表 1-2

实验者	n	$n(A)$	$f_n(A)$
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

试验表明：随着 n 的增大，$\frac{n(A)}{n}$ 在 $\frac{1}{2}$ 附近波动的幅度越来越小，逐渐稳定于 $\frac{1}{2}$，$\frac{1}{2}$ 这个值称为 $f_n(A)$ 的稳定值，通常把这个稳定值称为事件 A（出现正面）的概率. $f_n(A)$ 称为 n 次重复试验中 A 发生的频率，$n(A)$ 称为 n 次重复试验中 A 发生的频数. 下面给出频率的一般定义：

定义 1-1 设 E 为随机试验，A 为其中任一事件，$n(A)$ 为事件 A 在 n 次重复试验中发生的次数，则称比值 $\frac{n(A)}{n}$ 为 n 次试验中 A 发生的频率，记为

$$f_n(A) \triangleq \frac{n(A)}{n}, \tag{1-1}$$

其中 $n(A)$ 称为事件 A 在 n 次重复试验中发生的频数. 当 n 增大时，$f_n(A)$ 逐渐稳定于某一个确定值 $P(A)$，称 $P(A)$ 为频率的稳定值.

由频率的定义，不难看出频率 $f_n(A)$ 具有以下性质：

(1) 非负性：$0 \leqslant f_n(A) \leqslant 1$；

(2) 规范性：$f_n(\Omega)=1$；

(3) 可加性：若事件 A 与事件 B 互不相容，则

$$f_n(A \cup B)=f_n(A)+f_n(B), \tag{1-2}$$

进一步，若 $A_1, A_2, \cdots, A_m$ 两两互不相容，则

$$f_n\left(\bigcup_{i=1}^{m} A_i\right)=\sum_{i=1}^{m} f_n(A_i). \tag{1-3}$$

由掷硬币试验可见，用频率来刻画事件 A 发生的可能性大小比较直观，但它有随机波动的缺陷，即在一定条件下做重复试验，其结果可能是不一样的. 因此用频率的稳定值来刻画事

件 A 发生的可能性大小是比较恰当的.从而得概率的统计性定义如下:

定义 1-2 在不变的条件下做大量重复试验,称在重复试验中事件 A 发生的频率的稳定值 p 为事件 A 的概率,记为 $P(A)$.

尽管由频率的稳定值可得概率 $P(A)$,但我们不可能对每个事件都通过反复做试验以求得 $P(A)$.为此我们以频率的性质和频率的稳定值 $P(A)$ 为背景,采用抽象化方法给出概率 $P(A)$ 的一般定义.

2. 概率的公理化定义

定义 1-3 设 E 为随机试验,Ω 为它的样本空间,对 E 中的每一个事件 A 都赋予一个实数,记为 $P(A)$,且满足

(1) 非负性:$0\leqslant P(A)\leqslant 1$;

(2) 规范性:$P(\Omega)=1$;

(3) 可加性:若 $A_1,A_2,\cdots,A_n,\cdots$ 两两互不相容,有

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i). \tag{1-4}$$

则称 $P(A)$ 为事件 A 的概率.

由定义立得概率 $P(A)$ 的性质如下:

性质 1 $P(\varnothing)=0$.

性质 2 若 $A_1,A_2,\cdots,A_n$ 两两互不相容,则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i), \tag{1-5}$$

(1-5)式称为概率的有限可加性.

性质 3 若 A 的对立事件记为 $\overline{A}$,则

$$P(A)=1-P(\overline{A}). \tag{1-6}$$

证明 由于 $A\cup\overline{A}=\Omega$,且 $A\cap\overline{A}=\varnothing$,故由概率性质 2 及规范性得

$$1=P(\Omega)=P(A)+P(\overline{A}),$$

所以

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

性质 4 若 $A\subset B$,则

$$P(B-A)=P(B)-P(A), \tag{1-7}$$

且

$$P(A)\leqslant P(B).$$

证明 由于 $B=A\cup(B-A)$,而 $A\cap(B-A)=\varnothing$,所以由概率性质 2 得

$$P(B)=P[A\cup(B-A)]=P(A)+P(B-A),$$

故得

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

由非负性知,$P(B-A)\geqslant 0$,故 $P(B)-P(A)\geqslant 0$,即 $P(B)\geqslant P(A)$.

性质 5 对任意两个事件 A,B,有

$$P(A\cup B)=P(A)+P(B)-P(AB). \tag{1-8}$$

证明 由于 $A\cup B=A\cup(B-AB)$,而 $A\cap(B-AB)=\varnothing$,所以由概率性质 2 和性质 4 得

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

由概率性质 5 可推出任意有限个事件 $A_1, A_2, \cdots, A_n$ 的和事件概率的公式:

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)-\sum_{1\leqslant i<j\leqslant n} P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n} P(A_iA_jA_k)+\cdots +(-1)^{n-1}P(A_1\cdots A_n). \tag{1-9}$$

特别地,对于事件 A_1, A_2, A_3,有

$$P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3).$$

例 1-3 已知事件 A,B 满足 $P(AB)=P(\overline{A}\cap\overline{B})$ 且 $P(A)=p$,求 $P(B)$.

解法一 由概率性质知

$$\begin{aligned}P(B)&=P(A\cup B)-P(A)+P(AB) \quad (\text{概率性质 5})\\ &=1-P(\overline{A\cup B})-P(A)+P(AB) \quad (\text{概率性质 3})\\ &=1-P(\overline{A}\cap\overline{B})-P(A)+P(AB) \quad (\text{对偶原理})\\ &=1-P(A)=1-p. \quad (\text{已知条件})\end{aligned}$$

解法二 由于

$$\begin{aligned}P(AB)&=P(\overline{A}\cap\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)\\ &=1-P(A)-P(B)+P(AB),\end{aligned}$$

从而得

$$1-P(A)-P(B)=0,$$

即

$$P(B)=1-P(A)=1-p.$$

例 1-4 设对于事件 A,B,C,有 $P(A)=P(B)=P(C)=\dfrac{1}{4}$, $P(AC)=\dfrac{1}{6}$, $P(AB)=P(BC)=0$,求 A,B,C 至少出现一个的概率.

解 由于 $ABC\subset AB$,从而由概率性质 4 知,$P(ABC)\leqslant P(AB)=0$,又由概率公理化定义知 $P(ABC)\geqslant 0$,所以 $P(ABC)=0$. 从而由概率的加法公式得

$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)\\ &=\frac{1}{4}\times 3-\frac{1}{6}=\frac{7}{12}.\end{aligned}$$

§1.3 古典概型(等可能概型)

[引例 1-2] 掷一枚均匀硬币,则样本空间 $\Omega=\{H,T\}$(H:正面,T:反面);掷一颗骰子,则样本空间 $\Omega=\{1,2,3,4,5,6\}$.

引例 1-2 中这两个试验具有以下两个共同特征:

(1) 试验的样本空间所含基本事件的个数只有有限个;

(2) 每个基本事件出现的机会都是均等的.

我们称具有上述两个特征的概型为古典概型(等可能概型).

定义 1-4 设 $\Omega=\{e_1, e_2, \cdots, e_n\}$,如果 $P(e_1)=P(e_2)=\cdots=P(e_n)$,则称这一概型为等可能概型. 因为它是概率论发展初期的主要研究对象,故通常称为古典概型.

对于古典概型,显然有 $P(e_i)=\dfrac{1}{n}(i=1,2,\cdots,n)$. 因此,若 A 为 Ω 中的事件,且

$A=\{e_{i1},e_{i2},\cdots,e_{ik}\}$，由于基本事件两两互不相容，则

$$P(A)=\sum_{j=1}^{k}P(e_{ij})=\frac{k}{n},$$

即

$$P(A)=\frac{A\text{ 中包含基本事件的个数}}{\Omega\text{ 中基本事件的总数}},$$

简记为

$$P(A)=\frac{n(A)}{n(\Omega)},\tag{1-10}$$

其中，$n(A)$表示 A 中包含基本事件的个数，$n(\Omega)$表示 Ω 中基本事件的总数.

例 1-5 将一枚均匀硬币连续掷三次，观察正、反面出现的情况，求：

(1) 样本空间 Ω；

(2) 恰有一次出现正面的概率；

(3) 至少有一次出现正面的概率.

解 由题意知：

(1) $\Omega=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}$.

(2) 设事件 A 表示"恰有一次出现正面"，则 $n(A)=C_3^1=3$，且 $n(\Omega)=2^3$，故

$$P(A)=\frac{n(A)}{n(\Omega)}=\frac{3}{8}.$$

(3) 设事件 B 表示"至少有一次出现正面"，则 $n(B)=C_3^1+C_3^2+C_3^3=7$，且 $n(\Omega)=8$，故

$$P(B)=\frac{n(B)}{n(\Omega)}=\frac{7}{8}.$$

或用概率性质 3 先求逆事件的概率，即设 $\overline{B}$ 表示 B 的对立事件(三次中没有一次出现正面)，则 $n(\overline{B})=C_3^3=1$，所以 $P(\overline{B})=\frac{1}{8}$，故

$$P(B)=1-P(\overline{B})=\frac{7}{8}.$$

例 1-6 一个盒子中有 6 只球，其中有 4 只白球、2 只红球，从中取球两次，每次取一只，求下列事件的概率：

A：取到的两只球均为白球；

B：取到的两只球同色；

C：取到的两只球至少有一只为白球.

解 考虑两种抽样方式：(1) 有放回抽样，即第一次取出一只球，观察其颜色后放回盒子中，搅匀后再取第二只球；(2) 无放回抽样，即第一次取到的一只球不再放回盒子中，将剩余的球搅匀后再取第二只球.

(1) 有放回抽样：第一次取球，盒中有 6 只球可取，第二次取球，仍有 6 只球可取，因此 $n(\Omega)=6\times6$.

① 对于事件 A，第一次只能从 4 只白球中任取一只，第二次也只能从 4 只白球中任取一只，因此 $n(A)=4\times4$，所以

$$P(A)=\frac{n(A)}{n(\Omega)}=\frac{4\times4}{6\times6}=\frac{4}{9}.$$