普通高等教育“十二五”规划教材

概率统计及其应用

主编 张俊丽

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

普通高等教育“十二五”规划教材

概率统计及其应用

主　编　张俊丽
副主编　高陈燕　马明远

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

图书在版编目（CIP）数据

概率统计及其应用：含习题集/张俊丽主编．—北京：北京理工大学出版社，2014.9
ISBN 978-7-5640-9630-4

Ⅰ.①概…　Ⅱ.①张…　Ⅲ.①概率统计-高等学校-教学参考资料　Ⅳ.①O211

中国版本图书馆 CIP 数据核字（2014）第 194526 号

出版发行 /北京理工大学出版社有限责任公司
社　　址 /北京市海淀区中关村南大街 5 号
邮　　编 /100081
电　　话 /（010）68914775（总编室）
　　　　　82562903（教材售后服务热线）
　　　　　68948351（其他图书服务热线）
网　　址 /http：//www.bitpress.com.cn
经　　销 /全国各地新华书店
印　　刷 /北京地质印刷厂
开　　本 /787 毫米×1092 毫米　1/16
印　　张 /14.5
字　　数 /309 千字
版　　次 /2014 年 9 月第 1 版　2014 年 9 月第 1 次印刷
两册总定价 /29.80 元

责任编辑 /赵　岩
文案编辑 /赵　岩
责任校对 /孟祥敬
责任印制 /马振武

前　言 PREFACE

《概率统计及其应用》是研究随机现象统计规律性的数学学科，是高等学校各专业的一门重要的基础理论课. 随着现代科学技术的发展，概率统计在自然科学、社会科学、工程技术、工农业生产等领域中得到了越来越广泛的应用. 因此，在我国高等院校的绝大多数专业的教学计划中，概率统计被列为必修课程或限定选修课程.

作为一门应用数学学科，《概率统计及其应用》不但具有数学所共有的特点：高度的抽象性、严密的逻辑性和广泛的应用性，而且具有更独特的思维方法. 为使初学者尽快熟悉这种独特的思维方法，更好地掌握概率统计的基本概念、基本理论、基本运算以及处理随机数据的基本思想和方法，提高运用概率统计方法分析解决实际问题的能力和创造性思维能力，我们编写了本书. 本书分为概率和统计两个部分，其中第 1 章到第 4 章主要介绍概率论的基本知识，第 5 章到第 8 章介绍了数理统计的基本理论及常用多元统计分析，各章均配有习题及答案. 本书还有配套同步练习活页，可作为学生同步作业来使用. 本书的编写过程中，高陈燕编写了第 1 章和第 8 章的内容；张俊丽编写了第 2 章、第 3 章和第 4 章的内容；马明远担任编写了第 5 章、第 6 章、第 7 章的内容. 本书在编写的过程中得到了学院领导、部门领导和同事的支持与鼓励，在此表示感谢!

限于编者水平，教材中若存在不妥之处，希望广大读者批评指正. 联系方式：zll319@qq.com

目 录 CONTENTS

第 1 章　概率论的基本概念……1

§1.1　随机事件与事件关系……1

§1.2　概率及其性质……5

§1.3　古典概型和几何概型……7

§1.4　条件概率、全概率公式和贝叶斯公式……10

第 2 章　随机变量及其概率分布……18

§2.1　随机变量及其分布函数……18

§2.2　离散型随机变量及其概率分布……22

§2.3　连续型随机变量及其概率分布……28

§2.4　随机变量函数的概率分布……37

第 3 章　二维随机变量及其概率分布……44

§3.1　二维随机变量及其分布函数……44

§3.2　二维离散型随机变量……45

§3.3　二维连续型随机变量……46

§3.4　边缘分布……49

§3.5　随机变量的独立性……52

第 4 章　随机变量的数字特征……57

§4.1　数学期望……57

§4.2　方差……61

§4.3　协方差与相关系数……64

§4.4　切比雪夫不等式及大数定律……67

§4.5　中心极限定理……72

第 5 章　数理统计的概念……78

§5.1　总体与样本……78

§5.2　统计量……80

§5.3　χ^2 分布……84

§5.4　t 分布……88

§5.5　F 分布……90

第 6 章　参数估计……97

§6.1　点估计量的求法……98

§6.2　参数的区间估计 …… 103
第 7 章　假设检验 …… 113
§7.1　假设检验的基本概念 …… 113
§7.2　单个正态总体的均值与方差的假设检验 …… 115
§7.3　两个正态总体均值差与方差比的假设检验 …… 118
第 8 章　统计分析方法介绍及实现 …… 123
§8.1　相关分析和回归分析 …… 123
§8.2　聚类分析 …… 133
§8.3　判别分析 …… 136
附表一　泊松分布表 …… 142
附表二　正态分布表 …… 144
附表三　χ^2 分布表 …… 145
附表四　t 分布表 …… 147
附表五　p 值表 …… 149

第1章 概率论的基本概念

本章是概率论部分的基本概念和基本知识，是学习以后各章所必不可少的.

2000多年前，人们就开始玩20方块游戏.生活在公元前63年到公元14年的古罗马皇帝奥古斯都·凯撒在他的一封信中曾写到："我一整天都在玩骰子."

虽然人们赌博的历史很长，但概率论诞生的历史却相当短，部分原因在于，古人认为随机事件的出现是上帝意志的体现，人们没有必要去寻找事件出现的规律.直到文艺复兴时期的数学家卡尔达诺（1501—1576）出现，他是一个医生、占星家，也是一个赌徒，他写了一本书，叫《游戏机遇的学说》，又名《大术》.他写到："把一个骰子掷三次，得到某一给定点数的可能性至少是50%."他还写到"用两个骰子掷出10的概率是0.5"，"两个骰子共有36个结果".同时他还认为一个人的运气能决定一个随机事件的结果.更有名的科学家伽利略（1564—1642）也对随机事件的规律性感兴趣，"为什么投三个骰子时，10和11出现的频率要比9和12大?"他采用列举的办法进行了证明.

而真正的概率论始于法国数学家费马（1601—1665）与帕斯卡（1623—1642）的通信.帕斯卡18岁时就发明了机械计算机并卖出好几台，他参加过历史上最有名的一个数学俱乐部，讨论各种新思想.而费马则通晓5种语言，同时和当时许多优秀的数学家通信.1654年，十分热衷赌博的法国贵族梅雷向帕斯卡提出了著名的赌金分配问题，这中间的许多结论都促进了概率论的最初发展.

§1.1 随机事件与事件关系

1.1.1 样本空间和随机事件

1. 样本空间

大千世界，所遇到的现象不外乎两类.一类是确定现象，另一类是随机发生的不确定现象.如在标准大气压下，水加热到100℃时沸腾，是确定会发生的现象.用石蛋孵出小鸡，是确定不可能发生的现象.而买彩票中不中奖，适当条件下种子发芽等这类不确定现象叫作随机现象.

定义1 使随机现象得以实现和对它观察的全过程称为随机试验，记作E.随机试验满足以下条件：

（1）试验的所有可能结果是已知的或是可以确定的；

（2）每次试验究竟会发生什么结果是事先无法预知的（试验之后才知道）.

定义2 随机试验的所有可能结果组成的集合称为样本空间，记为Ω.试验的每一个可能结果称为样本点，记为ω.

在具体问题中，给定样本空间是研究随机现象的第一步.

例 1 一盒中有十个完全相同的球，号码分别为 $1,2,3,\cdots,10$，从中任取一球，观察其标号，用 ω_i 表示标号为 i，$i=1,2,\cdots,10$，则样本空间 $\Omega=\{1,2,\cdots,10\}$.

例 2 在研究英文字母使用状况时，通常选用这样的样本空间：

$$\Omega=\{\text{空格},A,B,C,\cdots,X,Y,Z\}$$

例 1、例 2 讨论的样本空间只有有限个样本点，是比较简单的样本空间.

例 3 讨论某寻呼台在单位时间内收到的呼叫次数，可能结果一定是非负整数而且很难制定一个数为它的上界，这样，可以把样本空间取为 $\Omega=\{0,1,2,\cdots\}$.

这样的样本空间含有无穷个样本点，但这些样本点可以依照某种顺序排列起来，称它为**可列样本空间**.

例 4 讨论某地区的气温时，自然把样本空间取为 $\Omega=(-\infty,+\infty)$ 或 $\Omega=[a,b]$.

这样的样本空间含有无穷个样本点，它充满一个区间，称它为**无穷样本空间**.

从这些例子可以看出，随着问题的不同，样本空间可以相当简单，也可以相当复杂，在今后的讨论中，都认为样本空间是预先给定的，当然对于一个实际问题或一个随机现象，考虑问题的角度不同，样本空间可能选择不同.

例 5 掷骰子这个随机试验，若考虑出现的点数，则样本空间 $\Omega=\{1,2,3,4,5,6\}$；若考虑的是出现奇数点还是出现偶数点，则样本空间 $\Omega=\{\text{奇数、偶数}\}$.

由此说明，同一个随机试验可以有不同的样本空间.

在实际问题中，选择恰当的样本空间来研究随机现象是概率中值得研究的问题.

2. 随机事件

定义 3 样本空间中满足一定条件的子集称为随机事件，简称事件，用大写字母 $A,B,C,\cdots$ 表示.

只含有一个不可再分试验结果的随机事件的称为一个基本事件，即一个样本点所组成的集合$\{\omega\}$. 在试验中如果出现随机事件 A 中包含了某一个样本点 ω，则称作事件 A 发生，并记作 $\omega\in A$.

再看例 1 样本空间 $\Omega=\{1,2,3,\cdots,10\}$，下面研究这些问题.

$A=\{\text{球的标号为 }3\}$， $B=\{\text{球的标号为偶数}\}$， $C=\{\text{球的标号不大于 }5\}$

其中 A 为一个基本事件，而 B 与 C 则是多个样本点的集合，即随机事件.

例如，B 发生(出现)必须而且只须下列样本点之一发生：$2,4,6,8,10$，它由五个样本点组成.

同样地，C 发生必须而且只须下列样本点之一发生：$1,2,3,4,5$.

显然 A,B,C 都是 Ω 的子集，它们可以简单地表示为 $A=\{3\}$，$B=\{2,4,6,8,10\}$，$C=\{1,2,3,4,5\}$.

因为 Ω 是由所有样本点所组成，因此在一次试验中，必然要出现 Ω 中的某一样本点 $\omega\in\Omega$，也就是在试验中 Ω 必然要发生，今后用 Ω 表示一个必然事件，空集 $\varnothing$ 不包含任何样本点，且在每次试验中总不发生，所以用 $\varnothing$ 表示不可能事件.

1.1.2 事件的关系与运算

对于随机试验而言,它的样本空间 Ω 可以包含很多随机事件,概率论的任务之一就是研究随机事件的规律,通过对较简单事件规律的研究掌握更复杂事件的规律,为此需要研究事件之间的关系与运算.

若没有特殊说明,认为样本空间 Ω 是给定的,且定义了 Ω 中的一些事件,$A_1,A_2,\cdots,A_i$ $(i=1,2,\cdots)$. 由于随机事件是样本空间的子集,从而事件的关系与运算和集合的关系与运算相类似.

1. 事件的包含关系

定义 4 若事件 A 发生必然导致事件 B 发生,则称事件 B 包含了事件 A,或称事件 A 包含于事件 B,记作 $A\subset B$ 或 $B\supset A$.

可以给上述含义一个几何解释:设样本空间是一个正方体,A、B 是两个事件,"A 发生必然导致 B 发生"意味着属于 A 的样本点在 B 中出现,由此可见,事件 $A\subset B$ 的含义与集合论是一致的.

例 6 例 1 中事件 $A=\{$球的标号为 6$\}$,事件 A 就导致了事件 $B=\{$球的标号为偶数$\}$的发生,因为摸到标号为 6 的球意味着偶数的球出现了,所以 $A\subset B$.

2. 事件的相等

定义 5 设 $A,B\subset\Omega$,若 $A\subset B$,同时有 $B\subset A$,称 A 与 B 相等,记为 $A=B$. 易知相等的两个事件 A、B 总是同时发生或同时不发生,在同一样本空间中两个事件相等意味着它们含有相同的样本点.

例如,例 1 中事件 $A=\{$球的标号为偶数$\}$和事件 $B=\{2,4,6,8,10\}$为相等事件.

3. 和(并)事件与积(交)事件

定义 6 设 $A,B\subset\Omega$,事件"A 与 B 中至少有一个发生"称为 A 和 B 的和事件或并事件,记作 $A\cup B$. 见图 1-1(a).

显然,$A\cup B=$"A 或 B 发生",$A\cup\varnothing=A$,$A\cup\Omega=\Omega$,$A\cup A=A$.

若 $A\subset B$,则 $A\cup B=B$,$A\subset A\cup B$,$B\subset A\cup B$.

若 n 个事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生,则称为 n 个事件的和,记作 $A_1\cup A_2\cup\cdots\cup A_n$ 或 $\bigcup\limits_{i=1}^{n}A_i$.

例 7 设某种圆柱形产品,若底面直径和高度都合格,则该产品合格. 令 $A=\{$直径不合格$\}$,$B=\{$高度不合格$\}$,则 $A\cup B=\{$产品不合格$\}$.

定义 7 设 $A,B\subset\Omega$,"A 与 B 同时发生"这一事件称为 A 和 B 的积事件或交事件,记作 AB 或 $A\cap B$. 见图 1-1(b).

显然,$A\cap\varnothing=\varnothing$,$A\cap\Omega=A$,$A\cap A=A$,$A\cap B\subset A$,$A\cap B\subset B$.

若 $A\subset B$,则 $A\cap B=A$.

若 n 个事件 $A_1,A_2,\cdots,A_n$ 同时发生,则称为 n 个事件的积事件,记作 $A_1A_2\cdots A_n$ 或 $\bigcap\limits_{i=1}^{n}A_i$.

如例 7 中，若 $C=\{$直径合格$\}$，$D=\{$高度合格$\}$，则 $CD=\{$产品合格$\}$.

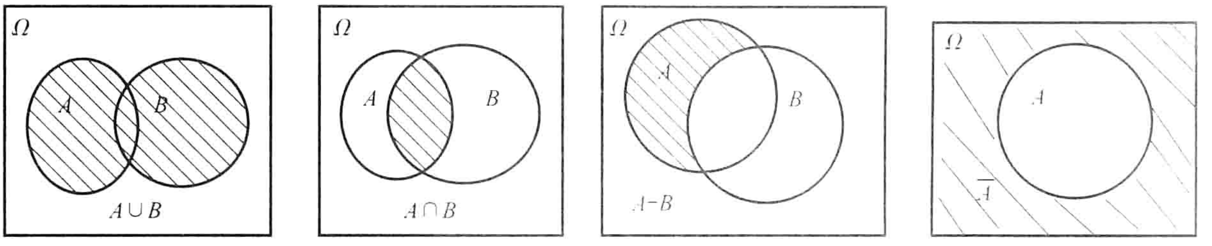
4. 差事件

定义 8 设 A、$B\subset\Omega$，"A 发生 B 不发生"这一事件称为 A 与 B 的差事件，记作 $A-B$. 见图 1-1(c).

如例 7 中，$A-B=\{$该产品的直径不合格，高度合格$\}$，显然有 $A-B=A-AB$，$A-\varnothing=A$.

5. 对立事件

定义 9 "$\Omega-A$"称为 A 的对立事件或称为 A 的逆事件，记作 $\overline{A}$. 见图 1-1(d).



(a) (b) (c) (d)

图 1-1

显然，$A\cup\overline{A}=\Omega$，$A\overline{A}=\varnothing$，$\overline{\Omega}=\varnothing$，$\overline{\varnothing}=\Omega$.

由此说明，在一次试验中 A 与 $\overline{A}$ 有且仅有一个发生，即不是 A 发生就是 $\overline{A}$ 发生.

显然 $\overline{\overline{A}}=A$，由此说明 A 与 $\overline{A}$ 互为逆事件.

例 8 设有 100 件产品，其中 5 件产品为次品，从中任取 50 件产品. 记 $A=\{$50 件产品中至少有一件次品$\}$. 则 $\overline{A}=\{$50 件产品中没有次品$\}=\{$50 件产品全是正品$\}$.

由此说明，若事件 A 比较复杂，往往它的对立事件比较简单，因此我们在研究复杂事件时，往往转化为研究它的对立事件.

6. 互不相容事件(互斥事件)

定义 10 若两个事件 A 与 B 不能同时发生，即 $AB=\varnothing$，称 A 与 B 为互不相容事件(或互斥事件).

注意：任意两个基本事件都是互斥的. 若 A、B 为互斥事件，则 A、B 不一定为对立事件. 但若 A、B 为对立事件，则 A、B 互斥.

推广：设 n 个事件 A_1、A_2、…、A_n 两两互斥，称 A_1、A_2、…、A_n 互斥(互不相容).

7. 事件的运算法则

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

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

(4) 对偶原则 $\overline{\bigcup\limits_{i=1}^{n}A_i}=\bigcap\limits_{i=1}^{n}\overline{A_i}$，$\overline{\bigcap\limits_{i=1}^{n}A_i}=\bigcup\limits_{i=1}^{n}\overline{A_i}$.

例 9 设 A,B,C 为 Ω 中的随机事件，试用 A,B,C 表示下列事件.

(1) A 与 B 发生而 C 不发生 $AB-C$ 或 $AB\bar{C}$

(2) A 发生，B 与 C 不发生 $A-B-C$ 或 $A\bar{B}\bar{C}$

(3) 恰有一个事件发生 $A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$

(4) 恰有两个事件发生 $AB\bar{C}\cup A\bar{B}C\cup AB\bar{C}$

(5) 三个事件都发生 ABC

(6) 至少有一个事件发生 $A\cup B\cup C$ 或(3)(4)(5)之并

(7) A,B,C 都不发生 $\bar{A}\bar{B}\bar{C}$

(8) A,B,C 不都发生 $\overline{ABC}$

(9) A,B,C 不多于一个发生 $\bar{A}\bar{B}\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$ 或 $\overline{AB\cup BC\cup CA}$

(10) A,B,C 不多于两个发生 $\overline{ABC}$

例 10 试验 E：袋中有三个球编号为 1,2,3，从中任意摸出一球，观察其号码，记 $A=\{$球的号码小于 3$\}$，$B=\{$球的号码为奇数$\}$，$C=\{$球的号码为 3$\}$.

试问：(1) E 的样本空间？(2) A 与 B，A 与 C，B 与 C 是否互不相容？(3) A,B,C 对立事件是什么？(4) A 与 B 的和事件、积事件、差事件各是什么？

解 设 ω_i 表示摸到球的号码为 i，$i=1,2,3$

(1) 则 E 的样本空间为 $\Omega=\{\omega_1,\omega_2,\omega_3\}$；

(2) $A=\{\omega_1,\omega_2\}$，$B=\{\omega_1,\omega_3\}$，$C=\{\omega_3\}$；

A 与 B，B 与 C 是相容的，A 与 C 互不相容；

(3) $\bar{A}=\{\omega_3\}$，$\bar{B}=\{\omega_2\}$，$\bar{C}=\{\omega_1,\omega_2\}$；

(4) $A\cup B=\Omega$，$AB=\{\omega_1\}$，$A-B=\{\omega_2\}$.

§1.2 概率及其性质

1.2.1 概率的概念

对于随机试验中的随机事件，在一次试验中是否发生，虽然不能预先知道，但是它们在一次试验中发生的可能性是有大小之分的. 比如掷一枚均匀的硬币，那么随机事件 A(正面朝上)和随机事件 B(正面朝下)发生的可能性是一样的(都为 0.5). 又如袋中有 8 个白球，2 个黑球，从中任取一球. 当然取到白球的可能性要大于取到黑球的可能性. 一般地，对于任何一个随机事件都可以找到一个数值与之对应，该数值作为此事件发生的可能性大小的度量.

历史上有人做过掷硬币的试验：

表 1-1 掷硬币试验

实验者	n	n_A	$f_n(A)$
蒲丰	4 040	2 048	0.507 0
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

从表 1-1 中可以看，不管什么人去抛，当试验次数逐渐增多时，$f_n(A)$ 总是在 0.5 附近摆

动而逐渐稳定在 0.5. 从这个例子可以看出，一个随机试验的随机事件 A 在 n 次试验中出现的频率 $f_n(A)$，它在一个常数附近摆动当试验的次数 n 逐渐增多时，而逐渐稳定于这个常数. 这个常数是客观存在的，“频率稳定性”的性质不断地为人类的实践活动所证实，它揭示了隐藏在随机现象中的规律性.

定义 1 随机事件 A 发生的可能性大小的度量(数值)，即频率的稳定值，称为 A 发生的概率，记为 $P(A)$.

对于一个随机试验来说，它发生可能性大小的度量是自身决定的，并且是客观存在的. 概率是随机事件发生可能性大小的度量是自身的属性.

1.2.2 概率的公理化定义

设随机试验的样本空间为 Ω，对于任意事件 A，有且只有一个实数 $P(A)$ 与之对应，满足如下公理：

公理 1 (非负性)$P(A)\geqslant 0$.

公理 2 (规范性)$P(\Omega)=1$.

公理 3 (完全可加性)对任意一列两两互斥事件 $A_1,A_2,\cdots$，有 $P\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=\sum\limits_{n=1}^{\infty}P(A_n)$.

则称 $P(A)$ 为事件 A 的概率.

公理化定义的广泛适用性是其突出的优点，但必须注意的是，这一定义并未解决如何确定概率的问题. 下面是一组由公理化定义可推出的概率的性质.

性质 1 $P(\varnothing)=0$，即不可能事件的概率为零.

注意：公理 2 和性质 1 反过来不一定成立：概率为 1 的事件不一定为必然事件，概率为 0 的事件不一定为不可能事件.

性质 2 对任一事件 A，均有 $P(\overline{A})=1-P(A)$.

性质 3 对任意有限个两两互斥事件 $A_1,A_2,\cdots,A_n$，有 $P\left(\bigcup\limits_{k=1}^{n}A_k\right)=\sum\limits_{k=1}^{n}P(A_k)$，即互斥事件的和事件的概率等于它们各自的概率和.

性质 4 对两个事件 A 和 B，若 $A\subset B$，则有 $P(B-A)=P(B)-P(A)$，$P(B)\geqslant P(A)$.

性质 5 (加法公式)对任意两个事件 A 和 B，有 $P(A\cup B)=P(A)+P(B)-P(AB)$.

性质 5 可以推广为：

$$\begin{aligned}P(A_1\cup A_2\cup A_3)=&P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)\\&-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3)\end{aligned}$$

例 1 设 A、B 互不相容，且 $P(A)=p$，$P(B)=q$. 试求：$P(A\cup B)$，$P(\overline{A}\cup B)$，$P(AB)$，$P(\overline{A}B)$，$P(\overline{A}\overline{B})$. 见图 1-2.

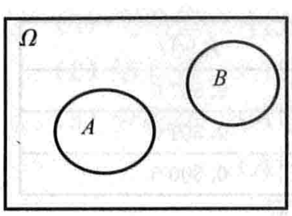


图 1-2

解 $P(A\cup B)=P(A)+P(B)=p+q$

$$P(\overline{A}\cup B)=P(\overline{A})=1-p\quad P(AB)=0$$

$$P(\overline{A}B)=P(B)=q\quad P(\overline{A}\overline{B})=1-P(A\cup B)=1-p-q$$

例 2 设 $P(A)=p$，$P(B)=q$，$P(A\cup B)=r$，求 $P(AB)$，$P(A\overline{B})$，$P(\overline{A}\cup\overline{B})$.

解 $P(AB)=P(A)+P(B)-P(A\cup B)=p+q-r$

$$P(A\bar{B})=P(A)-P(AB)=p-(p+q-r)=r-q$$
$$P(\bar{A}\cup\bar{B})=P(\overline{AB})=1-P(AB)=1-(p+q-r)=1-p-q+r$$

例3 设$P(A)=P(B)=P(C)=\frac{1}{8}$,$P(AB)=\frac{1}{16}$,$P(BC)=P(AC)=0$.求$A,B,C$至少有一个发生的概率.

解 A,B,C至少有一个发生的概率:

$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$

因为$ABC\subset BC$,所以$0\leqslant P(ABC)\leqslant P(BC)$,又因为$P(BC)=0$,所以$P(ABC)=0$

从而$P(A\cup B\cup C)=\frac{1}{8}+\frac{1}{8}+\frac{1}{8}-\frac{1}{16}=\frac{5}{16}$

问题思考:(配对问题)某人一次写了n封信,又写了n个信封,如果他任意将n张信纸装入n个信封中,问至少有一封信的信纸和信封是一致的概率是多少?

§1.3 古典概型和几何概型

1.3.1 古典概型

定义1 满足下列两个条件的这类现象称为古典概型:

(1) 样本空间中只有有限个样本点,即$\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$,其中$n$为样本点总数;

(2) 每个样本点$\omega_i(i=1,2,\cdots,n)$出现的可能性是相等的,并且每次试验有且仅有一个样本点发生.

古典概型事件概率的计算公式:

若事件A包含m个样本点,则事件A的概率定义为

$$P(A)=\frac{m}{n}=\frac{\text{事件}A\text{包含的样本点数}}{\text{样本点总数}}$$

古典概型在概率论中具有非常重要的地位.一方面它比较简单,既直观,又容易理解;另一方面它概括了许多实际内容,有很广泛的应用.

例如将一枚硬币连续掷两次就是这样的试验,也是古典概型,它有四个样本点:(正,正),(正,反),(反,正),(反,反),每个样本点出现的可能性都是$\frac{1}{4}$.

但将两枚硬币一起掷,这时试验的可能结果为(正,反),(反,反),(正,正),但它们出现的可能性却是不相同的,(正,反)出现的可能性为$\frac{2}{4}$,而其他的两个样本点出现的可能性为$\frac{1}{4}$.

例1 在盒子中有5个球(3个白球,2个黑球)从中任取两个,问取出的两个球都是白球的概率是多大?一白一黑的概率是多大?

分析:说明它属于古典概型,从5个球中任取2个,共有C_5^2种不同取法,可以将每一种取法作为一个样本点,则样本点总数C_5^2是有限的.由于摸球是随机的,因此样本点出现的可能性是相等的,因此这个问题是古典概型.

解 设$A=\{$取到的两个球都是白球$\}$,$B=\{$取到的两个球一白一黑$\}$,样本点总数为C_5^2.

事件 A 包含的样本点数为 C_3^2，$P(A)=\frac{C_3^2}{C_5^2}=\frac{3}{\frac{5\times 4}{2}}=\frac{3}{10}$

事件 B 包含的样本点数为 $C_3^1C_2^1$，$P(B)=\frac{C_3^1C_2^1}{C_5^2}=\frac{3\times 2}{\frac{5\times 4}{2}}=\frac{3}{5}$

由此例我们初步体会到解古典概型问题的两个要点：

(1)首先要判断问题是否属于古典概型，即要判断样本空间是否有限和样本点出现的等可能性；

(2)计算古典概型的关键是"记数"，这主要利用排列与组合的知识.

在古典概型时常利用摸球模型，因为古典概型中的大部分问题都能形象化地用摸球模型来描述. 若把黑球作为废品，白球看为正品，则这个模型就可以描述产品的抽样检查问题. 假如产品分为更多等级，例如一等品、二等品、三等品、等外品等，则可以用更多有多种颜色的摸球模型来描述.

例 2 在盒子中有十个相同的球，分别标为号码 1,2,3,…,10，从中任摸一球，求此球的号码为偶数的概率.

解法一 令 ω_i 表示摸到的球的号码为 i，$i=1,2,\cdots,10$，$\Omega=\{1,2,\cdots,10\}$，故样本点总数 $n=10$，令事件 $A=\{$所取球的号码为偶数$\}$，因而事件 A 含有 5 个样本点，因此有

$$P(A)=\frac{5}{10}=\frac{1}{2}$$

解法二 令事件 $A=\{$所取球的号码为偶数$\}$，则 $\overline{A}=\{$所取球的号码为奇数$\}$.

因而 $\Omega=\{A,\overline{A}\}$，$P(A)=\frac{1}{2}$

此例说明了在古典概型问题中，选取适当的样本空间，可使我们的解题变得简捷.

例 3 (分房问题)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($n\leqslant N$)，没有限制每间房住多少人，求下列事件的概率：

(1) $A=\{$指定的 n 个房间各有一人住$\}$；

(2) $B=\{$恰好有 n 个房间，其中各有一人住$\}$.

解 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.

(1) n 个人都分到指定的 n 间房中去住，保证每间房中各有一人住；

第一人有 n 分法，第二人有 $n-1$ 种分法，……最后一人只能分到剩下的一间房中去住，共有 $n(n-1)\cdots 2\cdot 1$ 种分法，即 A 含有 $n!$ 个样本点，因此有

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

(2) n 个人都分到的 n 间房中，保证每间只要一人，共有 $n!$ 种分法，而 n 间房未指定，故可以从 N 间房中任意选取，共有 C_N^n 种取法，故 B 包含了 C_N^n 种取法. 因此有

$$P(B)=\frac{C_N^n n!}{N^n}$$

例 4 (生日问题)某班级有 n 个人($n<365$)，问至少有两个人的生日在同一天的概率是多大？

解 假定一年按365天计算,将365天看成365个“房间”,那么问题就归结为分房问题.
令 $A=\{$至少有两个人的生日在同一天$\}$,$\overline{A}=\{n$ 个人的生日全不相同$\}$,

所以 $P(\overline{A})=\dfrac{C_N^n n!}{N^n}=\dfrac{N!}{N^n(N-n)!}(N=365)$

又因为 $P(A)+P(\overline{A})=1$

所以 $P(A)=1-P(\overline{A})=1-\dfrac{N!}{N^n(N-n)!}(N=365)$

这个例子就是历史上有名的“生日问题”,对于不同的一些 n 值,计算得相应的 $P(A)$ 如下表:

表 1-2

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

表中所列出的答案足以引起大家的惊奇,因为“一个班级中至少有两个人生日相同”,这个事件发生的概率并不如大多数人想象的那样小,而是足够大.从表中可以看出,当班级人数达到23时,就有半数以上的班级会发生这件事情,而当班级人数达到50人时,竟有97%的班级会发生上述事件,当然这里所讲的半数以上,有97%都是对概率而言的,只有在大数次的情况下(就要求班级数相当多),才可以理解为频率.这个例子告诉我们“直觉”并不可靠,从而更有力地说明了研究随机现象统计规律的重要性.

从上述几个例子可以看出,求解古典概型问题的关键是寻找样本点总数和事件 A 包含的样本点数,当正面求较困难时,可以转化求它的对立方面,要讲究一些技巧.

1.3.2 几何概型

定义2 当随机试验的样本空间是某一可度量的区域,并且任意一点落在度量(长度、面积与体积)相同的子区域内是等可能的,则事件 A 的概率定义为

$$P(A)=\frac{S_A}{S}=\frac{\text{构成事件 } A \text{ 的子区域的度量}}{\text{样本空间的度量}}$$

这种概率模型称为几何概型.

例5 (蒲丰投针问题)设平面上画有等距离为 a 的一簇平行线.取一枚长为 $l(l<a)$ 的针随意扔到平面上,求针与平行线相交的概率.

解 设 x 表示针的中心到最近一条平行线的距离,θ 表示针与此直线间的夹角,如图1-3(a)所示,则 (θ,x) 完全决定针所落的位置.针的所有可能的位置为

$$\Omega=\left\{(\theta,x)\mid 0\leqslant\theta\leqslant\pi,0\leqslant x\leqslant\frac{a}{2}\right\}$$

它可用 θ-x 平面上的一个矩形来表示,如图1-3(b)所示.针与平行线相交的充分必要条件是 $x\leqslant\dfrac{l}{2}\sin\theta$,即图1-3(b)中阴影部分,它的面积为

$$S_A=\int_0^\pi\frac{l}{2}\sin\theta\,\mathrm{d}\theta=l$$

因此,若把往平面上随意扔一枚针理解为 Ω 内的任一点等可能,且记针与任一平行线相交的事件为 A,则

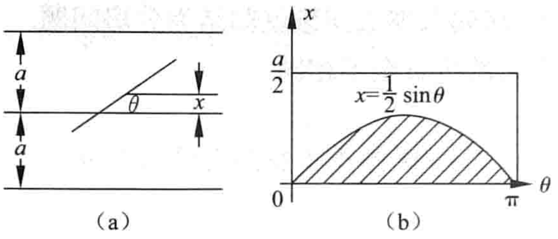


图 1-3

$$P(A)=\frac{S_A}{S}=\frac{2l}{\pi a} \tag{1}$$

由公式(1)可以利用投针试验计算 π 值，设随机投针 n 次，其中 k 次针线相交，当 n 充分大时，可用频率 $\frac{k}{n}$ 作为概率 $P(A)$ 的估计值，从而求得 π 的估计值为

$$\hat{\pi}=\frac{2ln}{ka} \tag{2}$$

根据公式(2)，历史上曾有一些学者做了随机投针试验，并得到了 π 的估计值.

§1.4 条件概率、全概率公式和贝叶斯公式

实际生活中出现的购买商品“以一赔十”、抓阄与次序是否有关、血液检查出阳性则患此病的可能性多大等问题，将在这节课给出解答.

前面讨论了事件和概率这两个概念，对于给定的一个随机试验，要求出一个指定的随机事件 A 的概率 $P(A)$，需要花很大的力气，现在继续深入讨论，设两个事件 A,B，则有加法公式

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

特别地，当 A,B 为互不相容的两个事件时，有 $P(A\cup B)=P(A)+P(B)$. 此时有 $P(A)$ 和 $P(B)$ 即可求得 $P(A\cup B)$. 但在一般情形下，为求得 $P(A\cup B)$ 还应该知道 $P(AB)$. 因而很自然要问，能不能通过 $P(A)$ 和 $P(B)$ 求得 $P(AB)$，先看一个简单的例子.

例 1 考虑有两个孩子的家庭，假定男女出生率一样，则两个孩子(依大小排列)的性别分别为(男，男)，(男，女)，(女，男)，(女，女)的可能性是一样的.

若记 $A=\{$随机抽取一个这样的家庭中有一男一女$\}$，则 $P(A)=\frac{1}{2}$. 但如果我们事先知道这个家庭至少有一个女孩，则上述事件的概率为 $P(A)=\frac{2}{3}$.

这两种情况下算出的概率不同，这很容易理解. 因为在第二种情况下我们多知道了一个条件. 记 $B=\{$这个家庭中至少有一个女孩$\}$，因此我们算得的概率是在已知事件 B 发生的条件下，事件 A 发生的概率，这个概率就是下面要介绍的条件概率.

1.4.1 条件概率和乘法公式

1. 条件概率的定义和乘法公式

定义 1 设 A 和 B 是两个事件，且 $P(B)>0$，即

$$\frac{P(AB)}{P(B)}$$

为在事件 B 已经发生的条件下，事件 A 发生的条件概率，记作 $P(A|B)$.

因此，例 1 中有 $P(A|B)=\frac{2}{3}$.

例2 某集体中有 N 个男人和 M 个女人，其中患色盲的男性 n 人，女性 m 人. 用 Ω 表示该集体，A 表示其中全体女性的集合，B 表示其中全体色盲者的集合. 如果从 Ω 中随意抽取一人，则这个人分别是女性、色盲者和同时既是女性又是色盲者的概率分别为：

$$P(A)=\frac{M}{M+N},\quad P(B)=\frac{m+n}{M+N},\quad P(AB)=\frac{m}{M+N}$$

如果限定只从女性中随机抽取一人(即事件 A 已发生)，那么这个女人为色盲者的(条件)概率为：

$$P(B\mid A)=\frac{m}{M}=\frac{P(AB)}{P(A)}$$

由条件概率公式可得，

$$P(AB)=P(A)P(B\mid A)=P(B)P(A\mid B) \tag{1}$$

称公式(1)为概率的乘法公式.

乘法公式可以推广到 n 个事件的情形：

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_n\mid A_1A_2\cdots A_{n-1}) \tag{2}$$

$$(P(A_n\mid A_1A_2\cdots A_{n-1})>0)$$

例3 (天气问题)甲、乙两市都位于长江下游，据100多年来的气象记录，知道在一年中，出现雨天的比例甲市占20%，乙市占18%，两地同时下雨占12%.

求：(1) 两市至少有一市是雨天的概率；

(2) 乙市出现雨天的条件下，甲市也出现雨天的概率；

(3) 甲市出现雨天的条件下，乙市也出现雨天的概率.

解 记事件 A={甲市出现雨天}，B={乙市出现雨天}，则根据题意有

$$P(A)=0.2,\quad P(B)=0.18,\quad P(AB)=0.12$$

(1) $P(A\cup B)=P(A)+P(B)-P(AB)=0.26$

(2) $P(A|B)=\dfrac{P(AB)}{P(B)}=0.67$

(3) $P(B|A)=\dfrac{P(AB)}{P(A)}=0.6$

例4 (零件出售“假一赔十”)已知商店出售某零件，每箱装这种零件100件，且包括4件次品. 假一赔十：顾客买一箱零件，如果随机取1件发现是次品，商店立刻用10件合格品取代其放入箱中. 某顾客在一个箱子中先后取了3件进行测试，求这3件都不是合格品的概率.

解 记 A_i={顾客在第 i 次取到次品}$(i=1,2,3)$，则有

$$P(A_1)=\frac{4}{100},\quad P(A_2\mid A_1)=\frac{3}{99+10}=\frac{3}{109},\quad P(A_3\mid A_1A_2)=\frac{2}{108+10}=\frac{2}{118}$$

根据乘法公式可知，顾客取出的3件都不是合格品的概率为

$$P(A_1A_2A_3)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)$$
$$=\frac{4}{100}\times\frac{3}{109}\times\frac{2}{118}=0.000\,02$$

2. 条件概率的性质

性质1 对于任意事件 A 和事件 B，有 $P(A|B)\geqslant 0$.