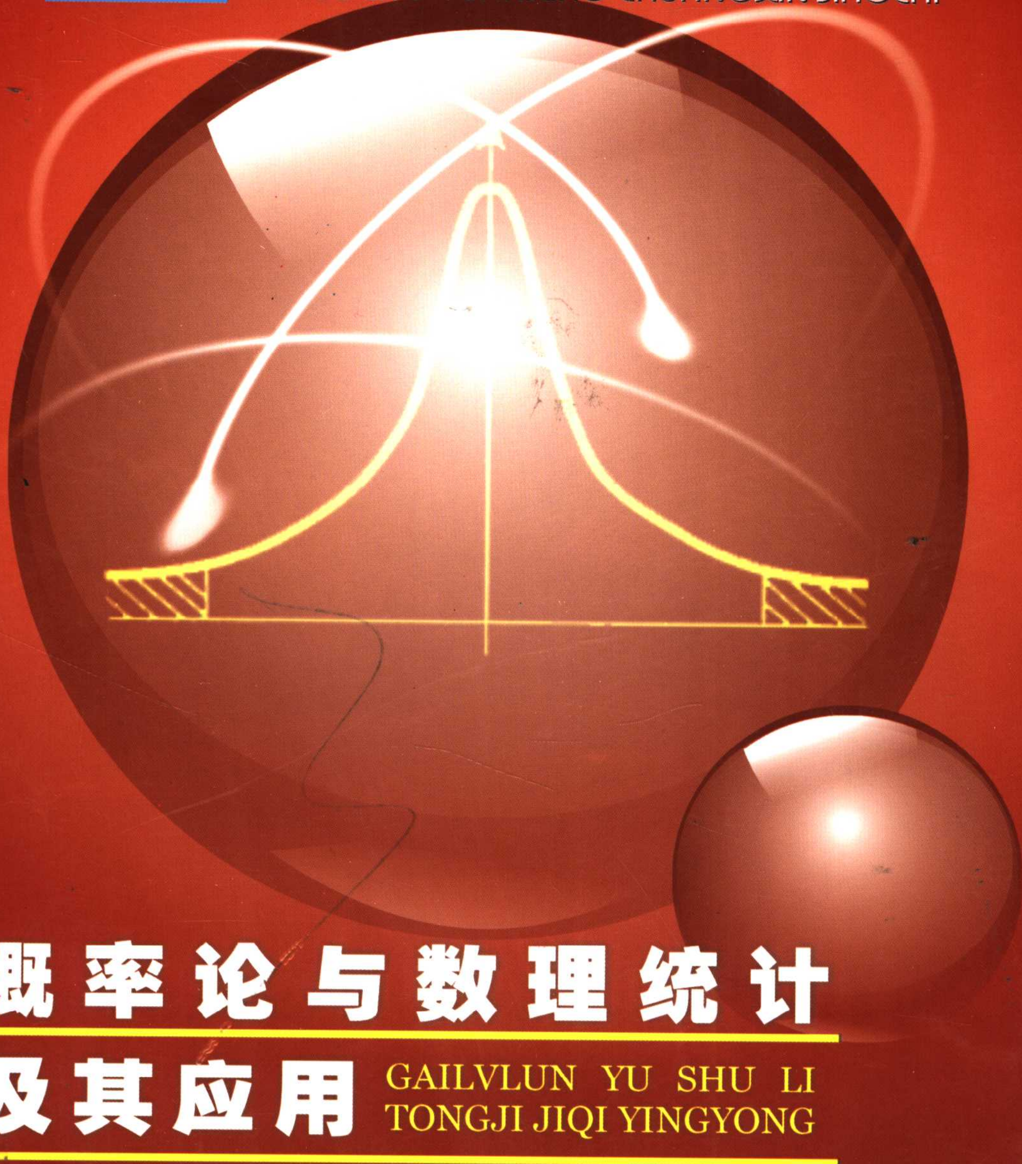
21世纪
高等院校创新教材
GAODENG YUANXIAO CHUANGXIN JIAOCAI



概率论与数理统计及其应用

GAILVLUN YU SHU LI TONGJI JIQI YINGYONG

熊德之 张志军 主编

科学出版社
www.sciencep.com

·21世纪高等院校创新教材·

概率论与数理统计及其应用

熊德之　张志军　主编

科学出版社

北京

内 容 简 介

本书是参照教育部高等学校数学与统计学教学指导委员会制订的《非数学类专业数学基础课程教学基本要求》编写而成. 全书共分十章，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、大数定律及中心极限定理、概率模型及其应用、数理统计的基本概念、参数估计、假设检验、方差分析与回归分析、SPSS 统计软件介绍与统计模型应用. 每章配有习题，书末附有习题答案.

本书特色鲜明：内容严谨，叙述详实，突出应用；针对概率论与数理统计课程实用性强的特点，增加了概率统计模型和统计软件介绍的内容，将数学建模思想、方法和实用软件工具融入一体. 本书可作为普通高等学校非数学类专业概率论与数理统计课程的教材，同时也可供教师、考研人员及工程技术人员参考.

图书在版编目（CIP）数据

概率论与数理统计及其应用/熊德之，张志军主编. –北京：科学出版社，2005

（21 世纪高等院校创新教材）

ISBN 7-03-016516-0

I.概… II.①熊…②张… III.①概率论–高等学校–教材②数理统计–高等学校–教材 IV.O21

中国版本图书馆 CIP 数据核字（2005）第 139328 号

责任编辑：冯贵层

责任印制：高 嵘／封面设计：李梦佳

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

武汉大学出版社印刷总厂印刷

科学出版社发行 各地新华书店经销

*

2005 年 12 月第 一 版 开本：B5（720×1000）

2005 年 12 月第一次印刷 印张：17 1/2

印数：1~5 000 字数：342 000

定价：25.00 元

（如有印装质量问题，我社负责调换）

前　言

概率论与数理统计是高等学校的一门重要基础课程，也是应用性极强的一门学科，它在理学、工学、农学、医学、经济学、管理学、军事学和体育科学等学科领域具有广泛的应用. 针对这门课程的特点，我们编写本书的基本思路是：着重概率统计思想和方法的阐述，着重学生应用能力的培养. 本书较详细地叙述了概率论与数理统计中主要概念和方法产生的背景及思想，注重理论联系实际，突出解决问题的思路，详尽介绍各种概率统计模型的掌握与应用；注重将数学建模思想和方法有机地融入到教材中，以期提高学生利用概率统计思想方法建立数学模型和利用统计分析软件解决实际问题的能力.

教材改革是教学改革的重要内容之一. 面向21世纪的概率统计教材一方面要在培养和提高学生应用能力上有较大突破，另一方面要在培养学生的数学素质，增强学生学习数学的兴趣上作探索. 一年一度的全国大学生数学建模竞赛已成为高等学校学生参赛人数最多、影响最大的一项课外科技活动，这就要求我们在数学基础课程的教学中介绍数学建模的思想和方法，为学生参赛作必要的准备. 本书也是省级重点教学研究项目《将数学建模思想和方法融入概率统计课程教学中的研究》课题的成果之一. 我们参照教育部高等学校数学与统计学教学指导委员会制订的《非数学类专业数学基础课程教学基本要求》，并按照“十五”国家级规划教材及教育部面向21世纪课程教材规划的要求，以及学生参加全国大学生数学建模竞赛的需要，集多年教学之经验，在备课教案和讲义的基础上，编写了这本教材. 为适应不同的教学对象和不同专业类别的教学需要，我们将有些内容打“*”号以便在教学中进行取舍.

本书由熊德之、张志军主编，负责拟定教材的编写大纲，并对全书进行修改、统稿和校审，以及组织在教学中试用. 罗进参加了部分内容的修改和校审工作. 各章的具体编写人员如下：第一章至第三章，郭光耀；第四章，罗进；第五章，刘任河；第六章至第八章，熊德之；第九章、第十章的第1至3节，严国义；第十章第4至5节及习题，张志军.

我们深知在众多概率论与数理统计经典教材面前，要写出一本体系结构新颖、特色鲜明且受到欢迎的教材是十分困难的. 在当今改革与探索的年代，我们的教材

也算是一家之言.这本教材的编写仅仅是我们工作的开始,我们将不断探索,为概率统计课程的教学改革尽一份力量.

由于编者水平有限,书中难免有缺点和错误,欢迎广大专家、同行和读者批评指正.

编　者

2005 年 7 月

目　录

第一章　随机事件与概率…………………………………………………… 1

§1　随机现象与随机试验 ……………………………………………… 1

§2　样本空间、随机事件…………………………………………………… 2

§3　概率的定义 ……………………………………………………………… 4

§4　古典概型 ………………………………………………………………… 8

§5　条件概率………………………………………………………………… 10

§6　独立性…………………………………………………………………… 15

习题一 ………………………………………………………………………… 20

第二章　随机变量及其分布 ……………………………………………… 24

§1　随机变量………………………………………………………………… 24

§2　离散型随机变量及其分布律………………………………………… 25

§3　随机变量的分布函数………………………………………………… 29

§4　连续型随机变量及其概率密度……………………………………… 32

§5　二维随机变量………………………………………………………… 39

§6　边缘分布及条件分布………………………………………………… 43

§7　相互独立的随机变量………………………………………………… 50

§8　随机变量的函数的分布……………………………………………… 53

习题二 ………………………………………………………………………… 61

第三章　随机变量的数字特征 …………………………………………… 66

§1　数学期望………………………………………………………………… 66

§2　方差……………………………………………………………………… 74

§3　矩………………………………………………………………………… 80

§4　协方差与相关系数…………………………………………………… 82

习题三 ………………………………………………………………………… 85

第四章　大数定律及中心极限定理 ……………………………………… 89

§1　大数定律………………………………………………………………… 89

§2　中心极限定理…………………………………………………………… 91

习题四 ………………………………………………………………………… 95

第五章　概率模型及其应用 ……………………………………………… 97

§1　概率基本模型概述…………………………………………………… 97

§2　随机模型………………………………………………………………… 98

习题五……………………………………………………………………………… 104
第六章　数理统计的基本概念……………………………………………………… 106
§1　样本与统计量 ………………………………………………………………… 106
*§2　直方图与经验分布函数 ……………………………………………………… 109
§3　抽样分布 ……………………………………………………………………… 112
习题六……………………………………………………………………………… 118
第七章　参数估计……………………………………………………………… 121
§1　点估计 ………………………………………………………………………… 121
§2　评价估计量的标准 …………………………………………………………… 128
§3　区间估计 ……………………………………………………………………… 130
习题七……………………………………………………………………………… 139
第八章　假设检验……………………………………………………………… 144
§1　假设检验的基本概念 ………………………………………………………… 144
§2　单个正态总体参数的假设检验 ……………………………………………… 148
§3　两个正态总体参数的假设检验 ……………………………………………… 151
*§4　非正态总体参数的假设检验 ………………………………………………… 153
§5　总体分布的假设检验 ………………………………………………………… 155
*§6　秩和检验 ……………………………………………………………………… 161
*§7　独立性检验 …………………………………………………………………… 162
习题八……………………………………………………………………………… 165
第九章　方差分析与回归分析………………………………………………… 170
§1　单因素试验的方差分析 ……………………………………………………… 170
§2　双因素试验的方差分析 ……………………………………………………… 180
§3　一元线性回归分析 …………………………………………………………… 188
*§4　多元线性回归模型简介 ……………………………………………………… 206
习题九……………………………………………………………………………… 209
***第十章　SPSS 统计软件介绍与统计模型应用** ……………………………… 213
§1　SPSS 11.0 for Windows 概述 ……………………………………………… 213
§2　数据文件的建立与整理 ……………………………………………………… 220
§3　SPSS 在描述统计与推断统计中的应用 …………………………………… 228
§4　数据、模型与统计应用………………………………………………………… 239
§5　统计模型实例 ………………………………………………………………… 243
习题十……………………………………………………………………………… 247
参考文献………………………………………………………………………… 250
习题答案………………………………………………………………………… 251
附表……………………………………………………………………………… 256

第一章　随机事件与概率

§1　随机现象与随机试验

1. 随机现象

概率论与数理统计是研究和揭示随机现象的统计规律性的一门数学学科. 现在首先介绍什么是随机现象,什么是统计规律性.

自然界和社会上发生的现象是多种多样的. 有一类现象,在一定条件下必然发生,例如,在没有外力作用的条件下,作匀速直线运动的物体必然继续作匀速直线运动,又如向上抛一石子必然下落,等等. 这类现象称为**确定性现象**. 但是,在自然界和社会上也还存在着另一类现象,例如,在相同条件下抛同一枚硬币,其结果可能是正面朝上,也可能是反面朝上,并且在每次抛掷之前无法肯定抛掷的结果是什么;用同一仪器多次测量同一物体的重量,由于受诸如测量仪器受大气影响、观察者生理上或心理上的变化等等偶然因素的影响,所得结果彼此总是略有差异. 这类现象,在一定的条件下,可能出现这样的结果,也可能出现那样的结果,而在试验或观察之前不能预知确切的结果. 但人们经过长期实践并深入研究之后,发现这类现象在大量重复试验或观察下,它的结果却呈现出某种规律性. 例如,多次重复抛一枚硬币得到正面朝上大致有一半,同一台仪器测量同一物体的重量,所得重量总在真实重量上下波动,等等. 这种在大量重复试验或观察中所呈现出的固有规律性,就是我们所说的**统计规律性**. 这种在个别试验中其结果呈现出不确定性,在大量重复试验中其结果又具有统计规律性的现象,称之为**随机现象**.

2. 随机试验

我们遇到过各种试验,在这里,我们把试验作为一个含义广泛的术语,它包括各种各样的科学实验,甚至对某一事物的某一特征的观察也认为是一种试验. 如果这个试验满足:

(1) 在相同条件下可以重复进行;

(2) 每次试验的可能结果不止一个,并且事先能够确定试验的所有可能结果;

(3) 在进行一次试验之前不能确定哪一个结果会出现.

我们称满足以上条件的试验为**随机试验**. 本书中所说的试验都是指随机试验. 下面举一些试验的例子.

E_1: 抛一枚硬币,观察正、反面出现的情况.

E_2：掷一颗骰子，观察出现的点数.

E_3：记录一天进入某超市的顾客数.

E_4：测试某种型号电视机的寿命.

E_5：测量某物理量（长度、直径等）的误差.

进行一次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 E_2 中，如果我们的目的是观察它朝上面的点数，其可能结果是：1 点，2 点，3 点，4 点，5 点，6 点；如果我们的目的是观察它朝上面点数的奇偶性，其可能的结果是：奇数点、偶数点两个.至于骰子落在桌面上哪个位置，朝哪个方向滚动等不在目的之列，不算作结果.

§2 样本空间、随机事件

1. 样本空间

对于随机试验，尽管在每次试验之前不能预知试验的结果，但试验的所有可能结果组成的集合是已知的，将随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**，记为 S. 样本空间的元素，即 E 的每个结果，称为**样本点**，记为 ω.

由于讨论的问题不同，其样本空间差别是很大的.下面介绍概率论与数理统计中通常讨论的一些情形.

例 1 将 2 颗骰子掷 1 次，观察其点数之和.其样本空间为 $\{2,3,4,\cdots,12\}$.

例 2 观察 1 小时中落在地球上某一区域的宇宙射线数.可能的结果一定是非负整数，而且很难指定一个数作为它的上界.这样，可以把样本空间取为 $\{0,1,2,\cdots\}$.

例 3 讨论某地区的气温时，我们可以把样本空间取为 $(-\infty,+\infty)$ 或 $[a,b]$.

实际上，随着问题的不同，其样本空间可以相当简单，如例 1；只有有限个样本点，也可能含有无穷多个点，如例 2，例 3.其中与例 2 类似的样本点数为可列个，像例 3 一样充满一个区间的是不可列无穷个.在实际问题中，有些样本空间可能比这些更复杂，但本书讨论问题的样本空间基本上只有以上三种情形.

2. 随机事件

有了样本空间的概念，就可定义随机事件.一般称试验 E 的样本空间 S 的子集为 E 的**随机事件**，简称**事件**.事件既可以看成样本空间的子集，也可以看成由样本点构成的集合.一般地，用字母 $A,B,C,\cdots$ 表示.在每次试验中，当且仅当这一集合中的一个样本点出现时，称这一**事件发生**.

不可能再分的事件称为**基本事件**.实际上，它是由一个样本点组成的单点集.由若干个基本事件组成的事件称为**复合事件**.特别地，样本空间 S 包含所有的样本

点，在每次试验中它总是发生的，称为**必然事件**. 空集$\varnothing$不包含任何样本点，它也可以作为样本空间的子集，它在每次试验中都不发生，称为**不可能事件**.

3. 事件间的关系与事件的运算

概率论的出现，是始于集合论之前的事情. 从本质上说，事件就是集合，事件间的关系与运算就是集合的关系与运算. 下面给出这些关系和运算在概率论中的提法，并根据“事件发生”的含义，给出它们在概率论中的含义.

(1) 若$A\subset B$，则称事件B包含事件A，指的是事件A发生必然导致事件B发生.

(2) 若$A\subset B$且$B\subset A$，即$A=B$，则称事件A与事件B**相等**.

(3) 事件$A\cup B=\{x\mid x\in A$ 或 $x\in B\}$称为事件A与事件B的**和事件**，当且仅当A,B至少有一个发生时，事件$A\cup B$发生.

类似地，称$\bigcup\limits_{k=1}^{n}A_k$为$n$个事件$A_1,A_2,\cdots,A_n$的和事件；称$\bigcup\limits_{k=1}^{\infty}A_k$为可列个事件$A_1,A_2,\cdots$的和事件.

(4) 事件$A\cap B=\{x\mid x\in A$ 且 $x\in B\}$称为事件A与事件B的**积事件**，当且仅当A,B同时发生时，事件$A\cap B$发生，$A\cap B$也记作AB.

类似地，称$\bigcap\limits_{k=1}^{n}A_k$为$n$个事件$A_1,A_2,\cdots,A_n$的积事件，称$\bigcap\limits_{k=1}^{\infty}A_k$为可列个事件$A_1,A_2,\cdots$的积事件.

(5) 事件$A-B=\{x\mid x\in A$ 且 $x\notin B\}$称为事件A与事件B的**差事件**. 当且仅当A发生，B不发生时，事件$A-B$发生.

(6) 若$A\cap B=\varnothing$，则称事件A与事件B是**互不相容的**，或**互斥的**. 这指的是事件A与事件B不能同时发生.

(7) 若$A\cup B=S$且$A\cap B=\varnothing$，则称事件A与事件B互为**逆事件**，又称事件A与事件B互为**对立事件**. 这指的是每次试验中，事件A,B中必有一个发生，且仅有一个发生. A的对立事件记为$\overline{A}$，且有$\overline{A}=S-A$，$A-B=A\cap\overline{B}$.

在进行事件运算时，要经常运用下述定律. 设A,B,C为事件，则有

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

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

摩根律　$\overline{A\cup B}=\overline{A}\cap\overline{B}$；$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

一般地，事件运算时的定律可推广到有限，如

$$\overline{A_1\cup A_2\cup\cdots\cup A_n}=\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_n};\quad \overline{A_1\cap A_2\cap\cdots\cap A_n}=\overline{A_1}\cup\overline{A_2}\cup\cdots\cup\overline{A_n}.$$

关于事件之间关系及运算与集合之间的关系及运算的类比，见下表1.1.

表 1.1　关于事件之间关系及运算与集合之间的关系及运算的类比

符号	概　率　论	集　合　论
S	样本空间或必然事件	空间(全集)
$\varnothing$	不可能事件	空集
ω	基本事件(样本点)	元素
A	事件 A	集合 A
$\overline{A}$	A 的对立事件	A 的余集
$A\subset B$	事件 A 发生必导致事件 B 发生	A 是 B 的子集
$A=B$	事件 A 与事件 B 相等	A 与 B 相等
$A\cup B$	事件 A 与 B 至少有一个发生	A 与 B 的并集
$A\cap B$	事件 A 与 B 同时发生	A 与 B 的交集
$A-B$	事件 A 发生而 B 不发生	A 与 B 的差集
$AB=\varnothing$	事件 A 和事件 B 互不相容	A 与 B 不相交

例 4　将下列事件用 A,B,C 表示出来：

(1) 只有 C 发生；

(2) 三个事件中至少有一个发生；

(3) 三个事件中不多于一个事件发生；

(4) 三个事件中不多于二个事件发生；

(5) 三个事件中恰好发生二个.

解　(1) $\overline{A}\,\overline{B}C$；

(2) $A\cup B\cup C$；

(3) $\overline{A}\,\overline{B}\,\overline{C}\cap\overline{A}\,\overline{B}C\cup A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}$；

(4) $\overline{A}\cup\overline{B}\cup\overline{C}$ 或 $\overline{A\cap B\cap C}$；

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

例 5　在理学院的学生中任选一名学生，其中理学院由数学系与物理系组成. 令事件 A 表示被选学生是女生，事件 B 表示该生是三年级学生，事件 C 表示该生是数学系学生.

(1) 叙述 $AB\overline{C}$ 的意义；

(2) 在什么条件下 $ABC=C$ 成立？

(3) 什么时候 $\overline{A}=B$ 成立？

解　(1) 该生是理学院物理系三年级女生.

(2) 在理学院的三年级数学系学生都是女生的条件下，$ABC=C$.

(3) 当理学院三年级学生都是男生，而其他年级都是女生时，$\overline{A}=B$.

§3　概率的定义

对于一个事件来说，我们不仅要掌握它在一次试验中是否发生，更进一步地，我们常常希望知道它在一次试验中发生的可能性究竟有多大. 为此，我们给出历史

上概率的二种最重要的定义及概率的基本性质.

1. 概率的统计定义

一个随机事件在某次试验中由于受许多无法控制的随机因素影响,不能断言它是否发生.若在相同的条件下,进行了n次试验,则将在这n次试验中A发生的次数记为n_A,称为事件A发生的**频数**,比值$\frac{n_A}{n}$称为事件A发生的**频率**,并记为$f_n(A)$.

由于事件A发生的频率是它发生次数与试验次数之比,其大小表示A发生的频繁程度,频率大,就意味着A在一次试验中发生的可能性就大,反之亦然.因而,直观的想法是用频率来表示A在一次试验中发生的可能性的大小.

但在实际上,A在n次试验中发生的频率受许多偶然性因素的影响,其表现为$f_n(A)$在0与1之间随机波动,但这种随机波动当试验次数越小,其幅度越大,试验次数越大,其幅度越小,也即随着试验次数的增多,频率稳定在某一常数附近,这种“频率稳定性”即通常所说的统计规律性.

历史上通过“掷一枚硬币”的试验来观察“出现正面”这一事件发生的规律,下表1.2是试验结果记录.

表1.2 历史上“掷一枚硬币”的试验结果

实验者	投掷次数	出正面次数	频率
蒲丰	4 040	2 048	0.5069
皮尔逊	12 000	6 019	0.5016
皮尔逊	24 000	12 012	0.5005

又如考察英文中特定字母出现的频率,当观察字母的个数n较小时,频率有较大的随机波动,但当n增大时,频率呈现出稳定性,下面就是一份英文字母的频率统计表.

表1.3 英文字母的频率统计表

字母	频率	字母	频率
E	0.1268	F	0.0256
T	0.0978	M	0.0244
A	0.0788	W	0.0214
O	0.0776	Y	0.0202
I	0.0707	G	0.0187
N	0.0706	P	0.0186
S	0.0634	B	0.0156
R	0.0594	V	0.0102
H	0.0573	K	0.0060
L	0.0394	X	0.0016
D	0.0389	J	0.0010
U	0.0280	Q	0.0009
C	0.0268	Z	0.0006

大量实验证实，在多次重复试验中，同一事件发生的频率虽然并不相同，但却在一个固定的数值附近摆动，呈现出一定的稳定性，而且随着重复试验次数的增加，这种现象愈加显著. 频率所接近的这个固定数值，就可作为相应事件的概率. 随机事件 A 的概率记为 $P(A)$.

定义1 在大量重复试验中，如果一个事件 A 发生的频率稳定在某一常数 p 附近摆动，这个数 p 就称为 A 的**概率**，记为 $P(A)=p$.

上述定义称为概率的统计定义.

概率的统计定义从直观上给出了概率的定义，但它却有理论上和应用上的缺点. 从理论上说，频率为什么具有稳定性呢（本问题将在第四章大数定律给出理论上的证明）？在应用上，我们没有理由认为，试验 $n+1$ 次来计算频率总会比试验 n 次更准确，更逼近所求的概率. 因此，我们不知道 n 取多大才行，如果 n 要很大，我们不一定能保证每次试验的条件完全一样，并且从感觉上讲，概率的统计定义也不像一种很严格的数学定义. 因而，要讨论概率的更加严格的数学化定义，即由前苏联数学家柯尔莫哥洛夫在1933 年给出的公理化定义. 由于其中涉及较多的实变函数等方面的知识，将其简化如下.

2. 概率的公理化定义

定义2 设 E 是随机试验，S 是它的样本空间，对于 E 的每一事件 A 赋予一个实数，记为 $P(A)$，称为事件 A 的**概率**（实际上，$P(A)$ 是一个定义域为集合、值域为 $[0,1]$ 的集合函数，简称集函数），并且集函数 $P(\cdot)$ 满足下列条件：

(1) **非负性**：对于每一个事件 A，有 $P(A)\geqslant 0$；

(2) **规范性**：对于必然事件 S，有 $P(S)=1$；

(3) **可列可加性**：设 $A_1,A_2,\cdots$ 是两两互不相容的事件，即对于 $i\neq j$，$A_iA_j=\varnothing$，$i,j=1,2,\cdots$，则有

$$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots,$$

即

$$P\Big(\bigcup_{k=1}^{\infty}A_k\Big)=\sum_{k=1}^{\infty}P(A_k). \tag{1.3.1}$$

下面，由概率的公理化定义，得到概率的一些重要性质.

性质1 $P(\varnothing)=0$. (1.3.2)

证 令 $A_n=\varnothing\,(n=1,2,\cdots)$，则 $\bigcup\limits_{n=1}^{\infty}A_n=\varnothing$，且 $A_iA_j=\varnothing$，$i\neq j$，$i,j=1,2,\cdots$，由可列可加性知

$$P(\varnothing)=P\Big(\bigcup_{n=1}^{\infty}A_n\Big)=\sum_{n=1}^{\infty}P(A_n)=\sum_{n=1}^{\infty}P(\varnothing).$$

得

$$P(\varnothing)=0.$$

性质2（有限可加性） 若 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件，则有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n). \quad (1.3.3)$$

(1.3.3)式称为概率的有限可加性.

证 令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$,则由可列可加性,得

$$P(A_1 \cup A_2 \cup \cdots \cup A_n)= P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{n} P(A_k) + 0$$
$$= P(A_1) + P(A_2) + \cdots + P(A_n).$$

性质3(单调性) 设 A,B 是两个事件,若 $A \subset B$,则有

$$P(A) \leqslant P(B). \quad (1.3.4)$$

证 因 $B=A\cup(B-A)$,$A\cap(B-A)=\varnothing$,则由性质2知

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

又由非负性知

$$P(B-A) \geqslant 0.$$

故

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

注意 式中等号即使在 A 是 B 的真子集的情况下也不可省.

性质4 设 A,B 是两个事件,若 $A \subset B$,则有

$$P(B-A) = P(B) - P(A). \quad (1.3.5)$$

证 由性质3证明可得.

性质5 对于任一事件 A,$P(A) \leqslant 1$.

证 因 $A \subset S$,由性质3易得.

性质6(逆事件的概率) 对于任一事件 A,有

$$P(\overline{A}) = 1 - P(A). \quad (1.3.6)$$

证 $A\cup\overline{A}=S$,$A\cap\overline{A}=\varnothing$,由性质2得

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

于是

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

性质7(加法公式) 对于任意两事件 A,B,有

$$P(A \cup B) = P(A) + P(B) - P(AB). \quad (1.3.7)$$

证 因 $A\cup B=A\cup(B-AB)$,且 $A\cap(B-AB)=\varnothing$,$AB\subset B$,故由(1.3.3)及(1.3.5)式,得

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

推论1 对于任意 n 个事件 $A_1,A_2,\cdots,A_n$,有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)+\cdots$$
$$+(-1)^{n-1}P(A_1A_2\cdots A_n). \quad (1.3.8)$$

用数学归纳法即可证得推论1.

推论2(次可加性) 对于任意 n 个事件 $A_1,A_2,\cdots,A_n$,有

$$P(\bigcup_{i=1}^{n} A_i) \leqslant \sum_{i=1}^{n} P(A_i). \quad (1.3.9)$$

§4 古典概型

以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，所以出现正面及反面的可能性都是$\frac{1}{2}$. 在概率论研究的初始阶段，主要讨论的随机事件都和上面例子一样具有两条性质：

(1) 试验的结果是有限的；

(2) 试验的每个结果是等可能的.

则对于任意事件A，对应的概率$P(A)$由下式计算

$$P(A)=\frac{\text{事件 } A \text{ 包含的基本事件数 } k}{S \text{ 中基本事件总数 } n}=\frac{k}{n}, \tag{1.4.1}$$

并把它称作古典概型.

在计算古典概型的概率时，主要是利用排列、组合来求数k与n. 要注意在计算k与n时是用排列还是组合，与次序是否有关等方面要一致.

例1 将二颗骰子掷一次，求它们点数之和为6点的概率.

解 点数之和为6点的事件记为A，将二颗骰子掷一次，如果考虑其点数之和，其样本空间为：

$$S_1=\{2,3,4,\cdots,12\};$$

如果考虑其点数组合，其样本空间为：

$$S_2=\{(1,1),(1,2),\cdots,(6,6)\} \quad (\text{共 } 36 \text{ 种情形}).$$

对于S_1出现的样本点，从直观上看，各个点数和概率相同显然是错误的. 而S_2中出现的各个样本点，从对称性可知其可能性相同. 故所求概率不能直接利用S_1而需利用S_2.

在S_2中基本事件总数为36种，而点数和为6点包含(1,5)，(2,4)，(3,3)，(4,2)，(5,1)等5种情形，故A的概率

$$P(A)=\frac{5}{36}.$$

当样本空间样本点不满足等可能性时，经常考虑转化为另一个满足等可能性的样本空间.

例2 设袋中有N件产品，其中有M件次品，从中取产品n次，每次随机地取一件. 考虑两种抽取方式：

(1) 先取一件产品后，观察它是否为正品，然后放回袋中，搅匀后再取一件，一直到取出n件为止，这种抽取方式叫做**有放回抽样**.

(2) 先取一件产品后，观察它是否为正品，然后从剩余的产品中再取一件，直到取出n件为止，这种抽取方式叫做**不放回抽样**.

试分别就上面二种抽样方式，求取到m件次品的概率.

解 (1) 先考虑有放回情况,此时样本空间S的样本点数为:第一次抽取时,有N种取法,第二次抽取时,仍有N种取法……如此下去,一共抽取n次,总共有N^n种等可能的样本点.记A_m为抽出的产品中有m件次品,则A_m所含样本点数为:首先在n次中选择m次,其选择方式有C_n^m种,在这m次取次品,其方法数为M^m,然后再在剩下的$n-m$次都取正品,有$(N-M)^{n-m}$种方法.故A_m的样本点数为$C_n^m M^m (N-M)^{n-m}$,则

$$P(A_m)=\frac{C_n^m M^m (N-M)^{n-m}}{N^n}=C_n^m\left(\frac{M}{N}\right)^m\left(1-\frac{M}{N}\right)^{n-m}.$$

(2) 再考虑不放回情况,其中$m\leqslant M, m\leqslant n$.先计算样本空间$S$的样本点数,从$N$件产品中任取$n$件,不讲次序,所有样本点的总数为$C_N^n$,记$B_m$为抽出的产品中有$m$件次品,则$B_m$所含样本点数为$C_M^m C_{N-M}^{n-m}$,可得

$$P(B_m)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n},\quad m\leqslant \min(n,M).$$

(1)中的分布称**二项分布**,(2)中的分布称**超几何分布**.

例3 从1,2,3,4,5,6六个数字中,等可能地有放回地连续抽取4个数字,试求下列事件的概率:

(1) $A=\{4$个数字完全不同$\}$;

(2) $B=\{4$个数字中不含1或5$\}$;

(3) $C=\{4$个数字中至少出现一次3$\}$.

解 从6个数字中有放回地抽取4次,因为每次抽取的事件数为6,故样本空间含基本事件数为6^4.

(1) A中元素个数是从6个数字中任意选取4个数字的排列数,因此有P_6^4个基本事件,则

$$P(A)=\frac{P_6^4}{6^4}=\frac{5}{18}.$$

(2) 令$B_1=\{4$个数字中不含1$\}$,$B_2=\{4$个数字中不含5$\}$,则$B=B_1\cup B_2$.又B_1中含基本事件数为5^4,B_2含基本事件数也为5^4,B_1B_2含基本事件数为4^4,则

$$\begin{aligned}P(B)&=P(B_1\cup B_2)=P(B_1)+P(B_2)-P(B_1B_2)\\&=\frac{5^4}{6^4}+\frac{5^4}{6^4}-\frac{4^4}{6^4}=\frac{497}{648}.\end{aligned}$$

(3) $\overline{C}=\{4$个数字中没有出现3$\}$,$\overline{C}$含基本事件数为5^4,则

$$P(C)=1-P(\overline{C})=1-\frac{5^4}{6^4}=\frac{671}{1296}.$$

例4 一袋中有a个红球,b个白球,从中任意地连续摸出k个球,每次摸出的球不放回袋中,试求最后一次摸到红球的概率.

解 将袋中每个球都看成是可以分辨的,抽球与次序有关,则样本空间的基本事件总数为P_{a+b}^k.设$A=\{$第k次摸到红球$\}$,则A所含基本事件数为:第一步,从a个

红球中任取一个红球，排在最后位置上，有a种方法；然后从剩下的$a+b-1$个球中取$k-1$个球任意放在$k-1$个位置，有P_{a+b-1}^{k-1}种方法. A含$a\times P_{a+b-1}^{k-1}$个基本事件，则

$$P(A)=\frac{a\times P_{a+b-1}^{k-1}}{P_{a+b}^{k}}=\frac{a}{a+b}.$$

从本例可以看出，摸到红球的概率与次序k无关，从而可以得到，在抽签中不论先后，每个人抽中的机会都是一样的，所以不必争先恐后.

进一步地，当采用有放回抽球时，每次抽到红球的概率也是$\frac{a}{a+b}$.

§5 条件概率

1. 条件概率的定义

在实际问题中，人们除了要考虑事件A的概率外，有时还需考虑在事件B发生的条件下，事件A发生的概率. 一般地说，两者的概率未必相同，为了区别起见，记后者为$P(A|B)$. 下面用一个例子说明这一点，并得到条件概率的定义.

例1 某年级有3个班，一班30人，二班35人，三班35人，在某次考试中，一班及格人数为28人，二班为30人，三班为32人，求总及格率及一班及格率.

解 该年级总人数为100人，总及格人数为90人，则总及格率为$\frac{28+30+32}{30+35+35}=\frac{9}{10}$，而一班的及格率为$\frac{28}{30}=\frac{14}{15}$.

实际上，一班的及格率也可以这样处理. 设A_i表示学生来自i班($i=1,2,3$)，B表示学生及格，则一班人数占总人数的百分率为$P(A_1)=\frac{30}{100}$，一班及格人数占总人数的百分率为$\frac{28}{100}$，记为$P(A_1B)$，一班及格率为$\frac{28}{30}=\frac{28/100}{30/100}$，记一班及格率为$P(B|A_1)$，则一般有

$$P(B|A_1)=\frac{P(A_1B)}{P(A_1)}.$$

类似地，对于一般古典概型问题，同上面讨论的一样，在加上某种限制条件后，对于任何事件A,B，当$P(A)\neq 0$时，总有

$$P(B|A)=\frac{P(AB)}{P(A)}.$$

故可以将上面关系式作为条件概率的定义.

定义1 设A,B是两个事件，且$P(A)>0$，则称

$$P(B|A)=\frac{P(AB)}{P(A)} \tag{1.5.1}$$

为在事件A发生的条件下事件B发生的**条件概率**.

不难验证，条件概率$P(\cdot|A)$符合概率定义中的三个条件，即