概率论与数理统计

邱望仁 主编 方成鸿 副主编

天津大学出版社
TIANJIN UNIVERSITY PRESS

概率论与数理统计

主　编　邱望仁
副主编　方成鸿

天津大学出版社
TIANJIN UNIVERSITY PRESS

内 容 提 要

本书是依据教育部对本课程的基本要求，以“应用为目的，理论是根基”为指导思想编写的. 主要内容包括概率论和数理统计，每章节后附有习题，书末附有参考答案. 本书由具有丰富教学经验的骨干教师编写，深入浅出，通俗易懂，便于自学.

全书共八章，第一章至第四章为概率论部分，包括概论率的基本概念、随机变量及其分布、多维随机变量及其分布、大数定律与中心极限定理；第五章至第八章为数理统计部分，包括样本及抽样分布、参数估计、假设检验、方差分析及回归分析.

本书可作为高等院校工学类、经济类、管理类本科“概率论与数理统计”教材，也可作为报考工学类、经济类、管理类研究生的复习参考书.

图书在版编目(CIP)数据

概率论与数理统计 / 邱望仁主编. — 天津：天津大学出版社，2015. 12

ISBN 978-7-5618-5493-8

Ⅰ. ①概… Ⅱ. ①邱… Ⅲ. ①概率论 - 高等学校 - 教材 ②数理统计 - 高等学校 - 教材 Ⅳ. ①O21

中国版本图书馆 CIP 数据核字(2015)第 321212 号

出版发行 天津大学出版社
地　　址 天津市卫津路 92 号天津大学内(邮编:300072)
电　　话 发行部:022-27403647
网　　址 publish. tju. edu. cn
印　　刷 廊坊市海涛印刷有限公司
经　　销 全国各地新华书店
开　　本 185mm × 260mm
印　　张 14.25
字　　数 362 千
版　　次 2015 年 12 月第 1 版
印　　次 2015 年 12 月第 1 次
印　　数 1 - 3 000
定　　价 36.00 元

前　言

当代巨变的人类文明给我们的生活带来了翻天覆地的变化,信息大爆炸也为教育工作者带来了新的挑战和机遇.作为高等院校的基础课教师,编写本书的课题组成员多年来一直在思考同一个问题:如何在新形势下,面对知识背景参差不齐的生源,找到一套有效的教育方式和方法,使培养的学生符合学校建设特色名校的理念,成为适应社会发展的新型人才.就“概率论与数理统计”这门课程而言,我们认为在教学过程中要面对并处理好以下几个问题.

首先,要搞清“概率论与数理统计”是一门什么样的课程?它是研究随机现象并找出其统计规律的一门学科.这门课程由概率论和数理统计两部分组成,是高等院校理工类、经管类的重要基础理论课程之一,属于数学最具特色的分支之一,是近代数学的重要组成部分.它具有自己独特的概念和方法,内容丰富,结论深刻.

其次,要明确“概率论与数理统计”在当代高等教育中应当处于什么样的地位?鉴于近年来该学科得到迅猛发展和广泛应用,它已成为一门独立的一级学科.随着计算机技术的发展,概率论与数理统计在自然科学、社会科学、工程技术、工农业生产等领域中得到了越来越深入的应用.与交叉学科得到飞速发展的同时,它也在向基础学科、工科学科渗透,与其他学科相结合发展衍生了一些边缘学科,这是概率论与数理统计发展的一个新趋势.因此,在我国高等学校的绝大多数专业的教学计划中,“概率论与数理统计”均被列为必修课程或限定选修课程.

最后,要在准确理解学校办学理念的前提下,设计出符合本校实际情况的教学内容和方法的载体.在人才资源越来越成为推动经济社会发展的战略性资源的时代背景下,特色名校的建设对于落实科学发展观、深化教育改革、合理配置教育资源有着重要意义,也有利于促进学生全面发展、个性成长,提高学校核心竞争力.我们课题组认为,一套好的教材是特色名校战略取得成功的重要条件,因为它是教学的媒介,是教师与学生沟通的桥梁,也是教学改革的重要组成部分.

正是基于上述三个方面的考虑,课题组成员共同努力,匠心编写了本教材,使它有着与众不同的特色.首先,本教材以“应用为目的,理论是根基”为指导思想编写,符合教育部对本课程的基本要求,立足于理论的理解,提高学生的解题

和知识的应用能力. 其次,教材的内容做了有针对性的修改. 限于教学课时,本教材着重于基本理论与概念的介绍,而没有把随机过程和时间序列等理论性较深的内容包括在内. 立足于陶瓷特色名校的建设,本教材最突出的特点是例子和插图的选择,在介绍基本理论和概念的过程中,所选择的例子大多与陶瓷生产和商业活动有着较密切的联系,希望学生在掌握概率论与数理统计基本理论的同时,对陶瓷行业有基本的认识,提高学生对此行业的理解. 另外,本教材的每章都介绍了一些历史人物及趣事,以加深学生对数学背景知识的理解,提高学习积极性.

本书共八章,课题组群策群力,共同确定了本书的编写指导思想与原则、例题选材与模式. 详细的编写工作中,邱望仁负责第一章的编写;第二章和附表中的表格由方成鸿主笔;第三、四、五章分别由汤文菊、肖楠、崔永琴负责完成;操群承担第六章和第七章的编写;第八章由徐明周博士主笔. 统稿工作由邱望仁和方成鸿完成,周永正教授与詹棠森教授等同行在本书编写过程中提供了很多有益的建议.

本书是课题组探索上述问题答案道路上的一块"原石",希冀在广大同行和读者的"雕琢"下成为一块"璞玉".

编　者

2015 年 12 月

目　　录

第一章　概率论的基本概念 …… (1)
第一节　样本空间与随机事件 …… (1)
第二节　频率与概率 …… (6)
第三节　古典概型与几何概率 …… (9)
第四节　条件概率 …… (16)
第五节　独立性 …… (22)
习题一 …… (24)
本章故事 …… (27)
第二章　随机变量及其分布 …… (29)
第一节　随机变量 …… (29)
第二节　离散型随机变量 …… (30)
第三节　分布函数 …… (36)
第四节　连续型随机变量 …… (38)
第五节　随机变量的函数的分布 …… (45)
第六节　随机变量的数字特征 …… (49)
习题二 …… (56)
本章故事 …… (60)
第三章　多维随机变量及其分布 …… (61)
第一节　二维随机变量 …… (61)
第二节　边缘分布与随机变量的独立性 …… (65)
第三节　条件分布 …… (71)
第四节　两个随机变量的函数的分布 …… (75)
第五节　多维随机变量的数字特征 …… (80)
习题三 …… (86)
本章故事 …… (88)
第四章　大数定律与中心极限定理 …… (90)
第一节　大数定律 …… (90)
第二节　中心极限定理 …… (93)
习题四 …… (96)
本章故事 …… (97)
第五章　样本及抽样分布 …… (101)
第一节　随机样本 …… (101)

第二节　直方图和箱线图 …………………………………………………… (102)
第三节　抽样分布 ………………………………………………………… (107)
习题五 ……………………………………………………………………… (115)
本章故事 …………………………………………………………………… (116)
第六章　参数估计 …………………………………………………… (117)
第一节　点估计 …………………………………………………………… (117)
第二节　估计量的评选标准 ……………………………………………… (124)
第三节　区间估计 ………………………………………………………… (127)
习题六 ……………………………………………………………………… (134)
第七章　假设检验 …………………………………………………… (137)
第一节　假设检验概述 …………………………………………………… (137)
第二节　单个正态总体的假设检验 ……………………………………… (143)
第三节　两个正态总体的假设检验 ……………………………………… (150)
第四节　分布拟合检验 …………………………………………………… (156)
习题七 ……………………………………………………………………… (159)
本章故事 …………………………………………………………………… (162)
第八章　方差分析及回归分析 ……………………………………… (166)
第一节　单因素试验的方差分析 ………………………………………… (166)
第二节　双因素试验的方差分析 ………………………………………… (174)
第三节　一元线性回归 …………………………………………………… (180)
第四节　多元线性回归 …………………………………………………… (190)
习题八 ……………………………………………………………………… (192)
本章故事 …………………………………………………………………… (194)
参考答案 ……………………………………………………………… (197)
附　表 ………………………………………………………………… (206)
附表一　几种常用的概率分布表 ………………………………………… (206)
附表二　标准正态分布函数值表 ………………………………………… (209)
附表三　泊松分布表 ……………………………………………………… (210)
附表四　t 分布上侧分位数表 …………………………………………… (215)
附表五　χ^2 分布上侧分位数表 ………………………………………… (216)
附表六　F 分布上侧分位数表 …………………………………………… (217)
参考文献 ……………………………………………………………… (222)

第一章　概率论的基本概念

在自然界和人类的社会实践中存在两类现象:确定性现象和随机性现象. 确定性现象是指在特定条件下,这种现象一定会发生或一定不会发生,如太阳从东方升起,动物一定会死亡,植物一定需要进行光合作用等. 而随机性现象则是指在相同的条件下,这种现象可能发生也可能不发生.

概率论是研究随机现象数量规律的数学分支,是一门研究事情发生的可能性的学问,其起源与赌博问题有关. 概率与统计的一些基本概念和简单方法早期主要用于赌博和人口统计模型. 随着人类社会实践的拓展,人们需要了解各种不确定现象中隐含的必然规律性,并用数学方法研究各种结果出现的可能性大小,从而产生了概率论,并使之逐步发展成一门严谨的学科. 概率与统计的方法日益渗透到各个领域,并广泛应用于自然科学、经济学、医学、金融保险甚至人文科学中.

本章主要是介绍一些概率论的基本概念,为后面的学习做准备.

第一节　样本空间与随机事件

一、随机试验

随机现象是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,而且事先不能预知确切的结果,但是相同条件下大量重复的随机试验却往往呈现出明显的数量规律. 这正是人们需要研究的,那么如何研究呢? 不能预知的原因会出现多种结果,因此需要分析各种结果的出现以及相互关联的规律性. 经过长期实践,人们发现在大量重复试验或观察下,随机现象的结果呈现出一定的规律性,这说明随机现象有偶然性的一面,也有必然性的一面,这种必然性表现为在大量重复试验或观察下所呈现出的固有规律性,这种规律性称为随机现象的统计规律性.

概率论与数理统计正是研究和揭示随机现象统计规律性的一门数学学科. 我们把对随机现象的观察认为是一种试验,具有以下特点的试验称为**随机试验**:

(1)(可重复性)可以在相同条件下重复进行;

(2)(可观测性)试验的结果可能不止一个且是可观测的,即所有的结果是明确的;

(3)(随机性)每次试验将要出现的结果是未知的,但总是上述可能结果中的一个.

我们是通过研究随机试验来研究随机现象的,今后所称的试验都是指随机试验. 下面举些随机试验的例子.

E_1:抛掷一枚硬币,观察其出现正面或反面的情况.

E_2:掷一枚骰子,观察出现的点数.

E_3:观察移动公司人工咨询台某座席一天内收到的客户咨询次数.

E_4:在一批灯泡中任意抽取一只,测试它的寿命.

E_5:检查裂纹釉烧制后的纹理是否符合要求(图 1-1-1).

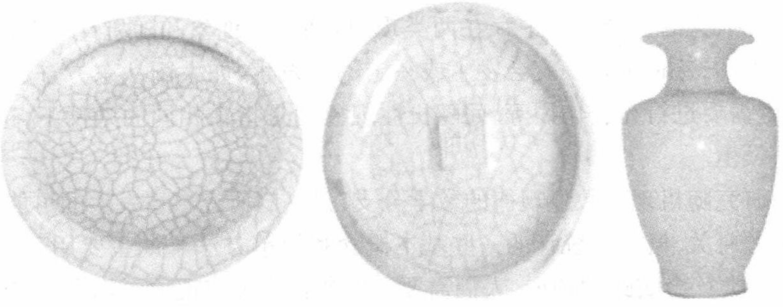

图 1-1-1

E_6:检查陶瓷的弯曲切口强度和弯曲断裂强度是否符合要求.

E_7:检查陶瓷活塞顶在受到一定的力后是否断裂(图 1-1-2).

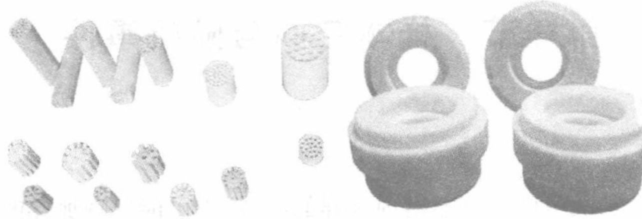

图 1-1-2

二、样本空间、随机事件

1. 样本空间

定义 1.1 随机试验的每一个可能结果称为**样本点**;一个随机试验的全体样本点构成的集合称为**样本空间**,记为 S.

根据上述举例,我们有:试验 E_1 的样本空间 $S_1 = \{正面,反面\}$,试验 E_2 的样本空间 $S_2 = \{1点,2点,3点,4点,5点,6点\}$,试验 E_3 的样本空间 $S_3 = \{n次 \mid n \in \mathbf{N}\}$,试验 E_4 的样本空间 $S_4 = \{t小时 \mid t \geqslant 0\}$,试验 E_5, E_6, E_7 的样本空间是{符合要求,不符合要求}.

如果用 1 表示出现正面,用 0 表示出现反面,试验 E_1 的样本空间 S_1 也可以表示为 $\{1,0\}$.

又设试验 E_8:观察贴花的两个瓷杯图案是否完全一样(图 1-1-3). 如果用 1 表示相同,用 0 表示不同,则 $S_8 = \{0,1\}$. 试验 E_9:抛掷两枚硬币,观察出现正面或反面的情况. 则 $S_9 = \{(正,正),(正,反),(反,正),(反,反)\}$.

以上表明,样本空间的元素是由试验的目的所确定的.

图 1-1-3

2. 随机事件

在随机试验中,除了关心各个样本点外,通常还关心满足指定特征的样本点在试验中是否发生. 例如对于试验 E_2,考察出现偶数点的情况,或者考察出现的点数大于 3 的情况.

设试验 E 的样本空间是 S,S 的子集称为 E 的**随机事件**,简称**事件**. 随机事件常用大写英文字母表示. 一次试验中,如果这个子集中的一个样本点出现,则称该**事件发生**.

由一个样本点组成的集合称为**基本事件**. 例如试验 E_2 有 6 个基本事件 $\{1\}$,$\{2\}$,$\{3\}$,$\{4\}$,$\{5\}$,$\{6\}$;试验 E_8 有两个基本事件{相同}和{不同}.

样本空间 S 是所有样本点构成的集合,因而在任何一次试验中,不管发生什么样的结果,都是 S 之中的某一种情况,即 S 必然发生,因此称 S 为**必然事件**. 经典集合中的空集 $\varnothing$ 不包括任何元素(样本点),因此在任何一次试验中都不可能有样本点属于它,也就是说空集 $\varnothing$ 永远不可能发生,因此我们把 $\varnothing$ 称为**不可能事件**.

例如:从含 3 件次品中的一套 32 头餐具中任取 10 件出来进行检查(图 1-1-4),在所选取的 10 件产品中,"次品多于 3 件"这一事件一定不会发生,为不可能事件;"次品不多于 3 件"这一事件一定会发生,为必然事件;而"次品有 3 件","次品有 2 件","次品有 1 件"等都是可能发生也可能不发生的事件,为随机事件.

图 1-1-4

3. 事件的运算与关系

根据前面的介绍,我们知道随机事件是样本空间的子集,其本质还是集合,而集合与集合之间存在着一定的关系,因此事件之间也存在着一定的关系,类似于集合的运算,也可以规定事件的运算,只不过其实际意义随研究的问题而变化. 下面给出这些关系和运算在概率论中的表述及其意义.

设试验 E 的样本空间是 S，而 $A,B,C,A_k(k=1,2,\cdots)$ 是 S 的子集. 下面先介绍事件的运算，然后再介绍事件的关系，接着介绍这些运算所满足的规律，最后用一些例子来解释说明它们在现实问题中的含义.

1）交（积事件）

事件 $A\cap B$ 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. 有时为了方便，$A\cap B$ 也写作 AB.

推广到 n 个事件 $A_1,A_2,\cdots,A_n$ 时，它们的积事件记为 $\bigcap\limits_{k=1}^{n}A_k$，也可记为 $A_1A_2\cdots A_n$.

两个事件的积可用图 1-1-5 表示.

两个事件的积可用集合的描述法表示为 $A\cap B=\{x|x\in A$ 且 $x\in B\}$.

2）并（和事件）

事件 $A\cup B$ 称为事件 A 与事件 B 的和事件，当且仅当 A、B 中至少有一个发生时，事件 $A\cup B$ 发生. 有时为了方便，$A\cup B$ 也写作 $A+B$. 两个事件的和可用图 1-1-6 表示.

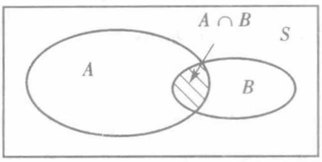

图 1-1-5

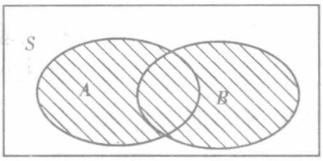

图 1-1-6

n 个事件 $A_1,A_2,\cdots,A_n$ 的和事件记为 $\bigcup\limits_{k=1}^{n}A_k$ 或 $A_1+A_2+\cdots+A_n$.

两个事件的和可用集合的描述法表示为 $A\cup B=\{x|x\in A$ 或 $x\in B\}$.

3）差

事件 $A-B$ 称为事件 A 与事件 B 的差事件，当且仅当 A 发生且 B 不发生时，事件 $A-B$ 发生.

两个事件的差可用集合的描述法表示为 $A-B=\{x|x\in A$ 且 $x\notin B\}$，如图 1-1-7 所示.

4）包含

如果 $B\subset A$，则称事件 A 包含事件 B，这时事件 B 发生必然导致事件 A 发生，如图 1-1-8 所示.

如果 $A\subset B$ 且 $B\subset A$，则称事件 A 与事件 B 相等，记为 $A=B$.

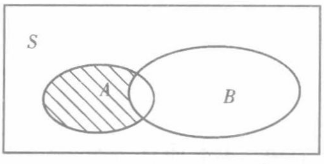

图 1-1-7

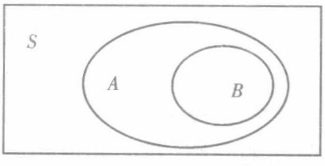

图 1-1-8

5）互不相容或互斥

如果 $A\cap B=\varnothing$，则称事件 A 与事件 B 是互不相容的，或互斥的. 互不相容的事件是不可

能同时发生的，基本事件是两两互不相容的. 两个事件的互不相容可用图 1-1-9 表示.

6）对立

如果 $A\cap B=\varnothing$ 且 $A\cup B=S$，则称事件 A 与事件 B 互为对立事件，或互为逆事件. 对每次试验来说，事件 A 与 B 必有一个发生，且仅有一个发生. A 的对立事件记为 $\bar{A}$，即 $\bar{A}=S-A$. 两个事件的对立可用图 1-1-10 表示（图中 $B=\bar{A}$，也有 $A=\bar{B}$）.

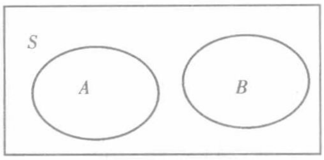

图 1-1-9

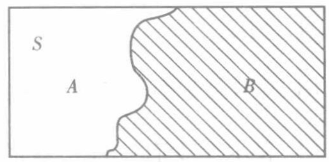

图 1-1-10

事件的运算律如下.

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：$(A\cup B)\cap C=AC\cup BC,(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$.

对偶律：$\overline{A\cup B}=\bar{A}\bar{B},\overline{AB}=\bar{A}\cup\bar{B}$；$\overline{\bigcup\limits_{k=1}^{n}A_k}=\bigcap\limits_{k=1}^{n}\overline{A_k},\overline{\bigcap\limits_{k=1}^{n}A_k}=\bigcup\limits_{k=1}^{n}\overline{A_k}$.

例 1-1-1　甲、乙、丙三个窑厂（图 1-1-11）参加某一工艺陶瓷的烧制比赛，因瓷窑设计工艺上的差异会导致瓷器在炼制过程中的技术参数相差比较大，最终会导致烧制的陶瓷制品有差别. 如果每个厂只烧制一次，记事件 A、B、C 分别表示“甲烧制成功”、“乙烧制成功”、“丙烧制成功”. 试用 A、B、C 三个事件表示下列各事件：

（1）丙厂没有烧制成功；

（2）三个厂中恰好有一个厂烧制成功；

（3）三个厂均未烧制成功；

（4）三个厂中至少两个厂烧制成功；

（5）三个厂中至多一个厂烧制成功.

解　（1）$\bar{C}$；（2）$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$；（3）$\bar{A}\bar{B}\bar{C}$；

（4）$AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup ABC$ 或 $AB\cup AC\cup BC$；

（5）$A\bar{B}\bar{C}\cup\bar{A}\bar{B}C\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}\bar{C}=\bar{A}\bar{B}\cup\bar{A}\bar{C}\cup\bar{B}\bar{C}$.

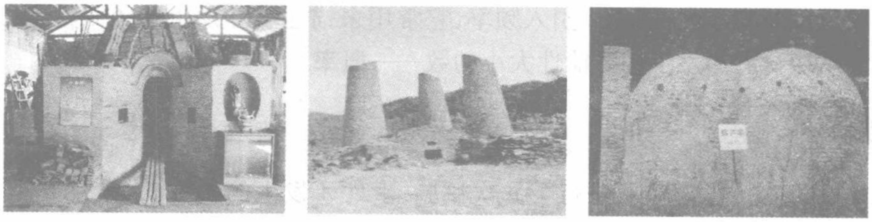

图 1-1-11

例 1-1-2　博物馆向某陶瓷工艺美术大师定做一件陶瓷，为确保能完成任务，此大师做

了三件，记 A_i 表示“第 i 件是合格产品”，$i=1,2,3$，记 B_j 表示“3 件中有 j 件是合格产品”，$j=0,1,2,3$. 用文字表述各组事件，并指出各组事件的关系：(1) $\overline{A_1}$ 与 B_3；(2) $A_1\overline{A_2}A_3$ 与 B_2；(3) $\bigcup\limits_{i=1}^{3}\overline{A_i}$ 与 $\overline{B_3}$；(4) $\bigcup\limits_{i=1}^{3}A_i$ 与 $\bigcup\limits_{i=1}^{3}B_i$.

解 (1) $\overline{A_1}$ 表示第一件不是合格产品，B_3 表示三件都是合格产品，它们是互斥事件；

(2) $A_1\overline{A_2}A_3$ 表示第二件不是合格产品且其余两件都是合格产品，B_2 表示有两件是合格产品，$A_1\overline{A_2}A_3\subset B_2$；

(3) $\bigcup\limits_{i=1}^{3}\overline{A_i}$ 表示至少有一件是不合格产品，$\overline{B_3}$ 表示三件不全是合格产品，$\bigcup\limits_{i=1}^{3}\overline{A_i}=\overline{B_3}$；

(4) $\bigcup\limits_{i=1}^{3}A_i$ 表示至少有一件是合格产品，$\bigcup\limits_{i=1}^{3}B_i$ 表示事件“有一件合格产品”、“有两件合格产品”与“有三件合格产品”至少有一个发生，$\bigcup\limits_{i=1}^{3}A_i=\bigcup\limits_{i=1}^{3}B_i$.

例 1-1-3 若 $S=\{1,2,3,4,5,6\}$，$A=\{1,3,5\}$，$B=\{1,2,3\}$，求：(1) $A\cup B$；(2) $A\cap B$；(3) $\bar{A}$；(4) $\bar{B}$；(5) $\overline{A\cup B}$；(6) $\overline{A\cap B}$；(7) $\bar{A}\cup\bar{B}$；(8) $\bar{A}\cap\bar{B}$.

解 (1) $A\cup B=\{1,2,3,5\}$；

(2) $A\cap B=\{1,3\}$；

(3) $\bar{A}=\{2,4,6\}$；

(4) $\bar{B}=\{4,5,6\}$；

(5) $\overline{A\cup B}=\{4,6\}$；

(6) $\overline{A\cap B}=\{2,4,5,6\}$；

(7) $\bar{A}\cup\bar{B}=\{2,4,5,6\}$；

(8) $\bar{A}\cap\bar{B}=\{4,6\}$.

第二节 频率与概率

虽然随机事件的发生有其偶然性，即在一次试验中它可能发生，也可能不发生；但在大量的重复试验中，往往还是会呈现一定的规律. 例如，投掷硬币的试验中，当试验次数很大时出现正反面的比率接近二分之一；投掷骰子的试验中，当投掷次数足够多时，每个点数出现的比率都接近六分之一. 对于这些随机事件，我们常常需要知道某个结果或某些结果在一次试验中发生的可能性有多大，即其内在规律，这些规律也是可以“度量”的. 为了方便说明这些概率统计的基本概念，下面先引入频率，它常用来描述事件发生的频繁程度，进而介绍刻画事件在一次试验中发生的可能性大小的数——**概率**.

一、频率

定义 1.2 在相同条件下，进行了 n 次试验，事件 A 发生的次数 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**，记为 $f_n(A)$.

试验表明随着试验次数 n 的增大，频率 $f_n(A)$ 呈现出稳定性，逐渐稳定于某个常数. 这

就是随机现象的统计规律.

由定义可以看出,在相同条件下,进行了 n 次试验,事件 A 发生的频率满足以下基本性质:

(1)(非负性)对任何事件 A,有 $0\leqslant f_n(A)\leqslant 1$;

(2)(规范性)对于必然事件 A,有 $f_n(A)=1$;

(3)(可列可加性)设 $A_1,A_2,\cdots$是两两互不相容的事件,那么

$$f_n(A_1\cup A_2\cup\cdots)=f_n(A_1)+f_n(A_2)+\cdots. \tag{1.2.1}$$

由于事件 A 发生的频率是它发生的次数与试验次数之比,其大小表示事件 A 发生的频繁程序. 频率越大表示事件 A 发生的越频繁,即意味着事件 A 在一次试验中发生的可能性就越大. 反过来,频率越小表示事件 A 发生的越不频繁,即意味着事件 A 在一次试验中发生的可能性就越小. 所以,人们常常用频率来表示事件 A 在一次试验中发生的可能性的大小.

这种想法已经得到了大量的试验证明,特别是早期概率论研究学者进行了深入的试验. 例如,表 1-2-1 和表 1-2-2 给出了抛硬币试验出现正面的频率结果. 表 1-2-1 中,将一枚硬币抛掷 5 次,50 次,500 次,各做 10 遍,得到的数据记录下来. 其中,n_A 表示事件 A"硬币出现正面"发生的频数,$f_n(A)$表示 A 发生的频率. 图 1-2-1 是表 1-2-1 中频率数据的变化趋势. 表 1-2-2 中的试验是由多个人做的,试验中仅记录每次抛掷试验中硬币出现正面的频数.

表 1-2-1　抛硬币试验结果(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

从上面的数据中可以看出,试验次数较少时,硬币出现正面的频率波动很大,然而随着试验次数的增加,硬币出现正面的频率波动在不断减弱. 这说明随着试验次数的增加,该事件发生的频率越来越稳定. 特别是当试验次数为 500 时,硬币出现正面的频率大致稳定在 0.5 附近. 而如表 1-2-2 所示,试验虽然是由不同人做的,但由于试验次数非常大,硬币出现正面的频率都是在 0.5 左右,其波动比表 1-2-1 中的还要小.

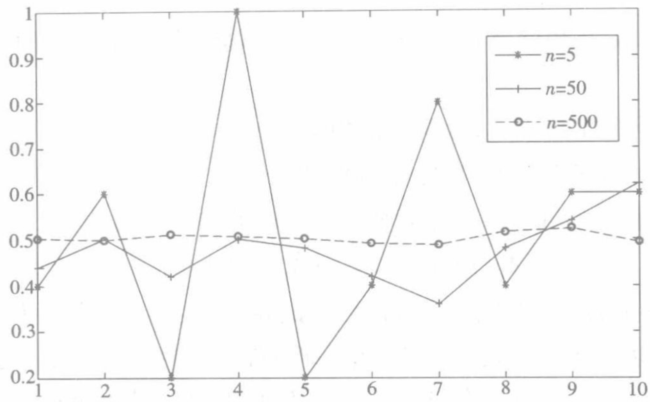

图 1-2-1

表 1-2-2　抛硬币试验结果(2)

实验者	n	n_A	$f_n(A)$
德・摩根	2 048	1 061	0.518 1
蒲丰	4 040	2 048	0.506 9
K. 皮尔逊	12 000	6 019	0.501 6
K. 皮尔逊	24 000	12 012	0.500 5

这说明当重复试验的次数 n 不断增大时,事件 A 发生的频率 $f_n(A)$ 呈现出稳定性,并会逐渐趋近于一个稳定的常数. 这种频率稳定性就是人们通常所说的统计规律. 也表明用频率来表示事件 A 在一次试验中发生的可能性的大小是可行的.

二、概率

尽管在大量试验中随机事件发生的频率会呈现稳定性,但人们不可能对每一个随机事件都做大量的试验,然后求得事件的频率,并用它来表示该随机事件发生的可能性的大小. 每次试验的结果虽然相差不大,但也存在细微的差别,不便于进行理论研究. 因此,概率论早期研究人员受到频率的稳定性及其能刻画事件发生可能性的特征的启发,给出了一个描述事件发生可能性大小的概念——概率.

定义 1.3(概率的定义)　设 E 是随机试验,S 是它的样本空间. 对于 E 的每一个事件 A 赋予一个实数,记为 $P(A)$,称为事件 A 发生的**概率**,若它满足以下三个条件:

(1)(非负性)对于任意事件 A,有 $0\leqslant P(A)\leqslant 1$;

(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,那么

$$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots. \tag{1.2.2}$$

事实上,这三点对应于前面所提到的事件 A 发生的频率所满足的基本性质,但频率只是试验统计意义,而概率是事件发生可能性的大小,具有理论意义.

根据概率的定义还可以推出以下一些重要的性质.

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

证明　由 $P(S)=P(S\cup\varnothing\cup\varnothing\cup\cdots)=P(S)+P(\varnothing)+P(\varnothing)+\cdots$,即得 $P(\varnothing)=0$.

(2)有限可加性:设 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件,那么 $P(\bigcup\limits_{i=1}^{n}A_i)=\sum\limits_{i=1}^{n}P(A_i)$;特别地,如果 $A\cap B=\varnothing$,那么 $P(A\cup B)=P(A)+P(B)$.

证明　令 $A_i=\varnothing,i=n+1,n+2,\cdots$,由概率的可列可加性及 $P(\varnothing)=0$ 即证.

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

证明　因为 $A\cup\bar{A}=S,A\cap\bar{A}=\varnothing$,由有限可加性及规范性,即得 $P(\bar{A})=1-P(A)$.

(4)设 $B\subset A$,则 $P(B)\leqslant P(A)$.

证明　因 $A=B\cup(A-B)$(图 1-1-8),且事件 B 与 $A-B$ 互不相容,故

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

又 $P(A-B)\geqslant0$,于是 $P(B)\leqslant P(A)$.

(5) $P(A-B)=P(A)-P(A\cap B)$;当 $B\subset A$ 时,$P(A-B)=P(A)-P(B)$.

证明　因 $A=(A-B)\cup(A\cap B)$(图 1-1-7),且 $(A-B)\cap(A\cap B)=\varnothing$,所以 $P(A)=P(A-B)+P(A\cap B)$,即 $P(A-B)=P(A)-P(A\cap B)$.

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

证明　因 $A\cup B=(A-A\cap B)\cup(B-A\cap B)\cup(A\cap B)$(图 1-1-6),且 $A-A\cap B$、$B-A\cap B$、$A\cap B$ 是两两互不相容的事件,所以

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

根据性质(5)知 $P(A-A\cap B)=P(A)-P(A\cap B)$,$P(B-A\cap B)=P(B)-P(A\cap B)$,于是

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

性质(6)可以推广到三个事件的情形:

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

例 1-2-1　已知 $P(\bar{A})=0.7$,$P(\bar{A}\cap B)=0.3$,$P(B)=0.4$,求 $P(A\cap B)$,$P(A-B)$.

解　因 $P(B)=P(A\cap B)+P(\bar{A}\cap B)$,所以 $P(A\cap B)=P(B)-P(\bar{A}\cap B)=0.1$,于是 $P(A-B)=P(A)-P(A\cap B)=1-P(\bar{A})-P(A\cap B)=0.2$.

第三节　古典概型与几何概率

一、排列与组合的相关内容

因为本节所介绍的古典概率模型需要应用到大量的排列组合知识,所以下面先将这些内容及相关公式罗列出来,以便于后面概率的计算.

1. 排列的定义

从 n 个不同元素中,任取 $m(m\leqslant n)$ 个元素(这里的被取元素各不相同)按照一定的顺序排成一列,叫作从 n 个不同元素中取出 m 个元素的一个**排列**. n 个不同元素全部取出的

排列,叫**全排列**.

根据排列的定义,两个排列相同,当且仅当两个排列的元素完全相同,且元素的排列顺序也相同. 例如,123 与 124 的元素不完全相同,它们是不同的排列;又如 123 与 132 虽然元素完全相同,但元素的排列顺序不同,它们也是不同的排列.

2. 排列数的定义

从 n 个不同元素中,任取 $m(m\leqslant n)$ 个元素的所有排列的个数叫作从 n 个元素中取出 m 个元素的排列数,用符号 P_n^m 表示.

例如,在 1,2,3 三个数中,任取两个元素能得到 12,13,21,23,31,32 共六种排列,即 $\mathrm{P}_3^2=6$.

3. 组合的定义

从 n 个不同元素中,任取 $m(m\leqslant n)$ 个元素(这里的被取元素各不相同)并成一组,叫作从 n 个不同元素中取出 m 个元素的一个**组合**.

例如,123 与 124 的元素不完全相同,它们是不同的组合;但对于 123 与 132,由于元素完全相同,尽管其元素的排列顺序不同,它们还是相同的组合,故只能算一个组合.

4. 组合数的定义

从 n 个不同元素中,任取 $m(m\leqslant n)$ 个元素的所有组合的个数叫作从 n 个元素中取出 m 个元素的**组合数**,用符号 C_n^m 表示.

例如,在 1,2,3 三个数中,任取两个元素只能得到 12,13,23 共三种组合,即 $\mathrm{C}_3^2=3$.

组合与排列的主要区别是组合是无序的,不考虑其中元素的先后次序;而排列是有序的,必须考虑元素的先后顺序.

5. 排列数公式

$$\mathrm{P}_n^m=n(n-1)(n-2)\cdots(n-m+1)=\frac{\mathrm{P}_n^n}{\mathrm{P}_{n-m}^{n-m}}=\frac{n!}{(n-m)!}.$$

其中,阶乘 $n!$ 为从自然数 1 到 n 的连乘积,在排列组合中也记为 P_n^n,且规定 $0!=1$.

6. 组合数公式

$$\mathrm{C}_n^m=\frac{\mathrm{P}_n^m}{\mathrm{P}_m^m}=\frac{n(n-1)(n-2)\cdots(n-m+1)}{m!}=\frac{n!}{m!\ (n-m)!}\ (n,m\in\mathbf{N}^*,m\leqslant n),$$

且有 $\mathrm{C}_n^m=\mathrm{C}_n^{n-m}$,$\mathrm{C}_{n+1}^m=\mathrm{C}_n^m+\mathrm{C}_n^{m-1}$,并规定 $\mathrm{C}_n^0=1$.

7. 几个常用公式

(1) $n\cdot n!=(n+1)!-n!$.

(2) $\dfrac{n}{(n+1)!}=\dfrac{1}{n!}-\dfrac{1}{(n+1)!}$.

(3) $\mathrm{C}_m^m+\mathrm{C}_{m+1}^m+\cdots+\mathrm{C}_n^m=\mathrm{C}_{n+1}^{m+1}$.

(4) $\mathrm{P}_m^m+\mathrm{P}_{m+1}^m+\cdots+\mathrm{P}_n^m=\mathrm{P}_m^m(\mathrm{C}_m^m+\mathrm{C}_{m+1}^m+\cdots+\mathrm{C}_n^m)=\mathrm{P}_m^m\cdot\mathrm{C}_{n+1}^{m+1}$.

8. 排列组合中常用的两个原理

(1) 分类计数原理:做一件事情,完成它可以有 n 类办法,在第一类办法中有 m_1 种不同的方法,在第二类办法中有 m_2 种不同的方法,……,在第 n 类办法中有 m_n 种不同的方法,