



普通高等教育“十二五”规划教材

概率论与数理统计

GAILVLUN YU SHULI TONGJI

主 编◎赵德平 陈仲堂
副主编◎宋介珠 徐厚生 李 莉 刘立士
主 审◎蔡 敏

概率论与数理统计

主 编	赵德平	陈仲堂		
副主编	宋介珠	徐厚生	李 莉	刘立士
参 编	赵恩良	孙丽华	沙 漠	邢双云
	隋 英	刘 丹	李卫国	
主 审	蔡 敏			

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

内 容 提 要

本书是为适应新的教学模式及现代科技对概率论与数理统计的需求,按照国家对非数学类本科生概率论与数理统计课程的基本要求编写的。辽宁省本科教改项目“新形势下概率与统计课程教学内容、教学方法与手段改革的研究与实践”研究的主要内容在教材编写中予以体现。

全书分为八章:随机事件及其概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、大数定律及中心极限定理、样本与抽样分布、参数估计、假设检验。各章配有习题,书末附有答案。全书切合实际教学情况,论述严谨,行文深入浅出,注重实用性。

本书即可作为高等院校理科、工科、经济、管理、生物等专业的教材使用,也可供科技人员和自学者参考。

版权专有 侵权必究

图书在版编目(CIP)数据

概率论与数理统计/赵德平,陈仲堂主编. —北京:北京理工大学出版社,2013.11
ISBN 978-7-5640-8467-7

I. ①概… II. ①赵… ②陈… 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(2013)第255250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/11.5

字 数/259千字

版 次/2013年11月第1版 2013年11月第1次印刷

定 价/26.80元

责任编辑/王俊洁

文案编辑/侯瑞娜

责任校对/周瑞红

责任印制/李志强

图书出现印装质量问题,请拨打售后服务热线,本社负责调换

编委会名单

主任委员：苏晓明 何希勤 徐送宁

副主任委员：赵 星 赵德平 聂 宏
孙丽媛 阎慧臻 石爱民
蔡 敏 宋岱才 霍满臣

编写说明

根据《教育部关于“十二五”普通高等教育本科教材建设的若干意见》(教高〔2011〕5号)精神和《辽宁省教育厅办公室关于组织开展“十二五”普通高等学校本科规划教材首批推荐遴选工作的通知》(辽教办发〔2011〕249号)的要求,沈阳工业大学、辽宁科技大学、辽宁石油化工大学、辽宁工业大学、大连交通大学、大连工业大学、沈阳航空航天大学、沈阳理工大学、沈阳建筑大学和沈阳工程学院等辽宁省内10所理工科院校理学院(数理系)发起组织了普通高等教育本科基础课高等数学(理工类)、高等数学(经管类)、概率论与数理统计、线性代数、工程数学、大学物理、大学物理实验、高等数学(英文·双语教材)、大学物理(英文·双语教材)等九门课程教材的编写工作。

为做好本套教材的编写工作,确保优质教材进课堂,辽宁省10所理工科院校的理学院院长(数理系主任)及基础课相关学科负责人组建了学科建设和教材编写专委会和编委会。专委会工作的目标是通过创新、融合,整合各院校优质教学教研资源,广泛吸收10所理工科院校在工科基础课课程教学理念、学科建设和体系搭建等方面的教学教研建设成果,按照当今最新的教材理念和立体化教材开发技术,通过不断的教材修订、立体化体系建设打造“工科基础课”教材品牌。

本套书力求结构严谨、逻辑清晰、叙述详细、通俗易懂。全书有较多的例题,便于读者自学,同时注意尽量多给出一些应用实例。

本书可供高等院校理工科类专业学生使用,也可供广大教师、工程技术人员参考。

辽宁省10所理工科院校理学院(数理系)
基础课学科建设和教材编写专委会和编委会
2013年6月6日

前 言

概率论与数理统计是一门既具特色又具有重要地位的工程数学课。它对提高学生的分析及处理不确定性现象的能力及运用概率统计方法解决实际问题的能力起着重要的作用。随着我国高等教育由“精英式”教育向“大众化”教育的转变,高等教育的模式较以往发生了较大的变化。加之现代科学技术的迅速发展及计算机的日益普及,对概率论与数理统计这门课提出了更高的要求,以往的教学模式及教学内容已不能完全适应现代科技的需求。如现在所用的教材重理论、轻实践,内容面面俱到,但学时却较少,很多讲不到。此外,现代科技要求我们增加一些应用环节,并要体现一些现代科技的内涵。为适应这一情况,我们联合沈阳建筑大学、沈阳理工大学、沈阳工业大学、辽宁科技大学等众多所高校,结合这些高校对这门课程的实际教学情况,融入陈仲堂教授负责的辽宁省本科教改项目“新形势下概率与统计课程教学内容、教学方法与手段改革的研究与实践”的研究成果,汇集这些高校教师多年的教学体会编写此书。

本书是非数学类本科生用概率论与数理统计课程的教材,其目的就是在教材中既要贯彻国家非数学类课程指导委员会制定的《概率论与数理统计课程的基本要求》,满足学生考研需求,又要适应工科院校本科生的特点,内容少而精,侧重实用性。

本书力求体现的特色如下:

(1)注意与中学数学教学内容及其他后续课的衔接,使其内容与中学的教学有机结合,而不是互相重叠或前后割裂。

(2)在保证教材内容的系统性及严谨性的基础上,注重在随机现象中数学概念的直观理解,注重数学观念的直观背景。

(3)以问题为驱动,由直观到抽象,由特殊到一般阐述内容。

(4)结合工科特点,注重理论知识在实际中的应用性,利用生活常识阐述概率与统计的理论,列举工程的实例来引领学生对内容的学习。

(5)致力于以近现代数学思想、观点和语言处理有关题材,采用现代统一的数学符号,并使其内容比传统的工科相应教材有较大的拓宽、充实、更新和提高,尽量体现现代科技的内涵。

全书论述严谨,行文深入浅出,注重实用性。希望学生能够通过本教材的学习,获得概率论与数理统计方面比较系统的知识,了解处理非确定现象一些常用的统计方法,为学生后续课程的学习及工作打下坚实的基础。

更重要的是,通过本教材的学习,使学生加深数学中处理随机现象的辩证统一思想的理解,并利用这一思想解决一些实际问题,全面提高学生的数学素质。

本书由沈阳建筑大学赵德平、陈仲堂主编。各章编写人员如下:赵恩良(沈阳建筑大学)、孙丽华(第一章,沈阳建筑大学)、宋介珠(第二章,辽宁科技大学)、李莉(第三章,沈阳工业大学)、沙漠(第四章,沈阳建筑大学)、邢双云(第五章,沈阳建筑大学)、刘立士(第六章,沈阳理工大学)、隋英(第七章,沈阳建筑大学)、陈仲堂(第八章,沈阳建筑大学)、刘丹(附表,沈阳建筑大学)。全书由赵德平、陈仲堂组织、构思及统纂,大连交通大学蔡敏教授给出审定意见,沈阳建筑大学徐厚生参与整体教材的编写组织工作,李卫国参与部分章节的编审工作。

由于编者水平有限,加之时间仓促,疏漏之处在所难免,恳请有关专家、同行及广大读者批评指正。

编 者

目 录

第一章 随机事件及其概率	1
§ 1 随机试验、样本空间及样本点	1
§ 2 随机事件及其运算	2
§ 3 概率	5
§ 4 条件概率与乘法原理	13
§ 5 随机事件的相互独立性	18
小结	22
习题一	22
第二章 随机变量及其分布	27
§ 1 随机变量	27
§ 2 随机变量的分布函数	28
§ 3 离散型随机变量及其分布律	30
§ 4 连续型随机变量及其概率密度	37
§ 5 随机变量的函数的分布	44
小结	48
习题二	48
第三章 多维随机变量及其分布	52
§ 1 二维随机变量及联合分布函数	52
§ 2 边缘分布	56
§ 3 条件分布	59
§ 4 随机变量的相互独立性	63
§ 5 随机变量的函数的分布	64
小结	68
习题三	70
第四章 随机变量的数字特征	72
§ 1 数学期望	72
§ 2 方差	78
§ 3 协方差及相关系数、矩	83
小结	86
习题四	87
第五章 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

小结	92
习题五	92
第六章 样本与抽样分布	94
§ 1 总体与简单随机样本	94
§ 2 抽样分布	96
小结	104
习题六	105
第七章 参数估计	106
§ 1 点估计	106
§ 2 估计量的评选标准	113
§ 3 区间估计	115
§ 4 正态总体的均值与方差的区间估计	118
小结	122
习题七	123
第八章 假设检验	125
§ 1 假设检验的基本概念	125
§ 2 正态总体均值的假设检验	129
§ 3 正态总体方差的假设检验	136
§ 4 分布拟合检验	140
小结	144
习题八	146
习题答案	149
附表	158
参考文献	173

第一章 随机事件及其概率

在自然界和社会生活中存在各种各样的现象,但归纳起来,无非是这样两类:一类是在一定条件下必然发生的,称之为确定性现象(也称必然现象).例如,早晨太阳必然从东方升起,同性电荷必相互排斥.另一类是指在一定的条件下,具有多种可能结果,而且事先不能预知哪种结果一定会出现或一定不出现,这类现象称之为随机现象(也称偶然现象).例如,抛一枚硬币,观察哪一面朝上,抛硬币前,我们不能确定正面朝上还是反面朝上;用同一门炮向同一目标射击,各次弹着点不尽相同,在一次射击之前不能预测弹着点的确切位置,等等.

人们经过长期实践及深入研究之后,发现随机现象虽然就每次试验或观察结果来说,具有不确定性,但是经过大量重复试验或观察,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;同一门炮射击同一目标弹着点按照一定规律分布,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,称之为随机现象的统计规律性.

概率论与数理统计就是研究随机现象统计规律性的一门学科.其理论与方法广泛应用于各个学科分支和各个生产部门.

§ 1 随机试验、样本空间及样本点

1.1.1 随机试验

为了研究随机现象的统计规律性,须进行各种观察或试验.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.一般说来,在概率论中,如果一个试验具备以下三个特点:

- (1)可以在相同的条件下重复进行;
- (2)每次试验的可能结果不止一个,且事先可明确试验的全部可能结果;
- (3)每次试验之前不能确定哪个结果一定会出现或一定不出现,

则称之为随机试验,简称试验,通常用字母 E 表示.我们是通过研究随机试验来研究随机现象的.

下面举一些试验的例子.

E_1 : 投掷一个篮球,观察球是否入筐;

E_2 : 掷一枚硬币两次,观察正面 H 、反面 T 出现的情况;

E_3 : 统计某车站在下午 1:00~2:00 之间的顾客数;

E_4 : 从一批灯泡中任取一只,测试它的使用寿命;

E_5 : 掷一颗骰子,观察出现的点数;

E_6 : 记录某出租车公司电话订车中心一天内接到订车电话的次数;

E_7 : 记录某射手射击半径为一米的圆盘靶弹着点的位置.

1.1.2 样本空间

对于随机试验, 尽管在每次试验前不能预知试验的结果, 但是试验的所有可能结果是已知的. 我们将随机试验 E 的所有可能结果组成的集合称为样本空间, 记为 S ; 样本空间的元素, 即 E 的每一个结果, 称为样本点, 记为 e .

例如, 上述随机试验的样本空间及样本点可表述如下:

E_1 : $S_1 = \{\text{投中, 没投中}\}$, 含两个样本点;

E_2 : $S_2 = \{HH, HT, TH, TT\}$, 含 4 个样本点;

E_3 : $S_3 = \{0, 1, 2, \dots\}$, 样本点为可数无穷多;

E_4 : $S_4 = \{t | t \geq 0\}$, 样本点为不可数无穷多;

E_5 : $S_5 = \{1, 2, \dots, 6\}$, 含 6 个样本点;

E_6 : $S_6 = \{0, 1, 2, 3, \dots\}$, 样本点为可数无穷多;

E_7 : $S_7 = \{(x, y) | x^2 + y^2 \leq 1\}$, 样本点为不可数无穷多.

§ 2 随机事件及其运算

1.2.1 随机事件

一般地, 我们把试验 E 的样本空间 S 的子集称为 E 的随机事件, 简称事件, 通常用英文大写字母 $A, B, C, \dots$ 表示. 在每次试验中, 当且仅当这一子集中的一个样本点出现时, 称这一事件发生.

特别地, 由一个样本点组成的单点集, 称为基本事件.

例如, 试验 E_1 中有两个基本事件: $\{\text{投中}\}$ 和 $\{\text{没投中}\}$; 试验 E_2 中有 4 个基本事件: $\{HH\}, \{HT\}, \{TH\}, \{TT\}$.

样本空间 S 包含所有的样本点, 它是 S 自身的子集, 在每次试验中它总是发生的, S 称为必然事件, 空集 $\emptyset$ 不包括任何样本点, 它也作为样本空间的子集, 它在每次试验中都不发生, $\emptyset$ 称为不可能事件.

例 1 掷一颗骰子, 观察出现的点数, 设 $A_i = \{i\} (i=1, 2, \dots, 6)$ 表示“每次掷一颗骰子出现的点数”, 是基本事件; 设 $B_1 = \{1, 3, 5\}$ 表示“掷一颗骰子时出现奇数点”, $B_2 = \{2, 4, 6\}$ 表示“掷一颗骰子时出现偶数点”; 设 $C_1 = \{1, 2, 3\}$ 表示“掷一颗骰子时点数是 1 或 2 或 3”; $C_2 = \{4, 5, 6\}$ 表示“掷一颗骰子时点数是 4 或 5 或 6”; $S = \{1, 2, \dots, 6\}$ 为必然事件.

1.2.2 随机事件的关系

事件是一个集合, 因而事件间的关系与运算自然按照集合论中集合之间的关系和运算来处理. 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, 给出它们在概率中的含义.

设试验 E 的样本空间为 S , 而 $A, B, A_k (k=1, 2, \dots)$ 是 S 的子集.

1. 事件的包含与相等关系

定义 1 若事件 A 发生必然导致事件 B 发生, 则称 A 包含于 B , 记为 $A \subset B$. 它的几何表示如图 1-1 所示. 若 $A \subset B$ 且 $B \subset A$, 则称 A 与 B 相等, 记为 $A=B$.

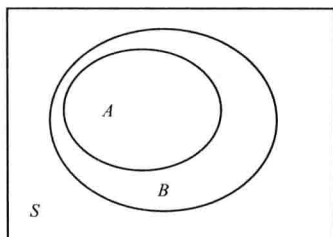


图 1-1

性质 1 (1) $\emptyset \subset A \subset S$; (2) $A \subset B, B \subset C \Rightarrow A \subset C$.

2. 事件的和

定义 2 事件 $A \cup B = \{x | x \in A \text{ 或 } x \in B\}$ 称为事件 A 与事件 B 的和事件, 当且仅当 A 与 B 中至少有一个发生时, $A \cup B$ 发生. $A \cup B$ 也记作 $A+B$. 它的几何表示如图 1-2 所示.

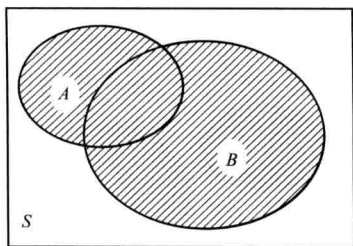


图 1-2

推广: 类似地, 我们称 $\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的和事件; 称 $\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \dots \cup A_n \cup \dots$ 为可列个事件的和事件.

3. 事件的积

定义 3 事件 $A \cap B = \{x | x \in A \text{ 且 } x \in B\}$ 称为事件 A 与事件 B 的积事件, 当且仅当 A 与 B 同时发生时, $A \cap B$ 发生. $A \cap B$ 也记作 AB . 它的几何表示如图 1-3 所示.

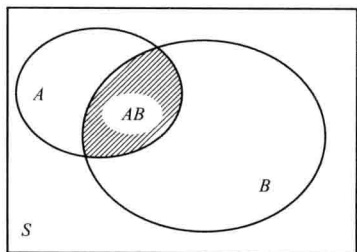


图 1-3

推广: 类似地, 我们称 $\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap \dots \cap A_n$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的积事件;

称 $\bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots$ 为可列个事件的积事件.

4. 事件的差

定义 4 事件 $A-B = \{x | x \in A \text{ 且 } x \notin B\}$ 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. 它的几何表示如图 1-4 所示.

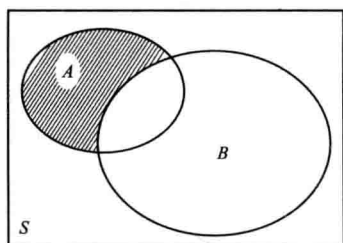


图 1-4

性质 2 若 $AB = \emptyset$, 则 $A-B = A$.

5. 互斥事件

定义 5 若 $A \cap B = \emptyset$, 称事件 A 与事件 B 互斥或互不相容, 此时事件 A 与事件 B 不能同时发生. 基本事件是两两互斥的. 它的几何表示如图 1-5 所示.

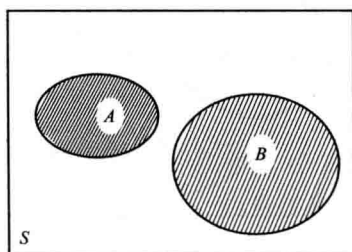


图 1-5

6. 对立事件

定义 6 若 $AB = \emptyset$ 且 $A \cup B = S$ 时, 称事件 A 与事件 B 互为对立事件或者互为逆事件. 它的几何表示如图 1-6 所示.

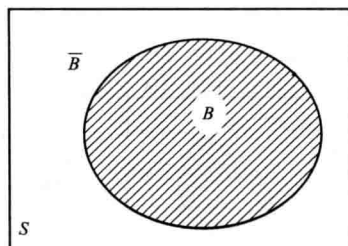


图 1-6

对于每次试验而言, 对立事件中有且仅有一个发生. A 的对立事件记为 $\bar{A}$. $\bar{A} = S - A$.

性质 3 (1) $A\bar{A} = \emptyset$, $A \cup \bar{A} = S$;

(2) $\overline{\bar{A}} = A$;

(3) 对立事件一定是互不相容事件, 但是互不相容事件未必是对立事件.

例 2 一个盒子中有 10 个完全相同的球, 分别标以号码 1, 2, ..., 10, 从中任取一球, 令

$A = \{\text{球的标号为偶数}\}$, $B = \{\text{球的标号} \leq 3\}$, 求: (1) $A \cup B$; (2) $A \cap B$; (3) $A - B$; (4) $\bar{A}$.

解 (1) $A \cup B = \{\text{球的标号为 } 1, 2, 3, 4, 6, 8, 10\}$;

(2) $A \cap B = \{\text{球的标号为 } 2\}$;

(3) $A - B = \{\text{球的标号为 } 4, 6, 8, 10\}$;

(4) $\bar{A} = \{\text{球的标号为奇数}\}$.

1.2.3 随机事件的运算规律

在进行事件运算时, 经常要用到下述定律. 设 A, B, C 为事件, 则有:

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

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

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

(4) 德·摩根律 $\overline{A \cup B} = \bar{A} \cap \bar{B}$; $\overline{A \cap B} = \bar{A} \cup \bar{B}$.

例 3 在射击比赛时, 一选手连续向目标射击三次, 令 A_i 表示“第 i 次射击, 击中目标” ($i=1, 2, 3$), 将下列事实用 A_i 表示出来:

(1) 只击中第一枪; (2) 只击中一枪; (3) 三枪均未中;

(4) 至少中一枪; (5) 最多中两枪; (6) 至少中两枪.

解 (1) “只击中第一枪”, 说明第二、三次射击都没有击中目标, 所以可表示为 $A_1 \bar{A}_2 \bar{A}_3$;

(2) “只击中一枪”未指明哪一枪击中, 因此是“只击中第一枪”“只击中第二枪”和“只击中第三枪”中的三个两两互斥的事件中的某一个发生, 所以可表示为 $A_1 \bar{A}_2 \bar{A}_3 \cup \bar{A}_1 A_2 \bar{A}_3 \cup \bar{A}_1 \bar{A}_2 A_3$;

(3) “三枪均未中”显然是 $\bar{A}_1 \bar{A}_2 \bar{A}_3$;

(4) “至少中一枪”为 $A_1 \cup A_2 \cup A_3$, 也可表示为 $\overline{\bar{A}_1 \bar{A}_2 \bar{A}_3} = A_1 \cup A_2 \cup A_3$;

(5) “最多中两枪”意味着未中三枪或者说 $\bar{A}_1, \bar{A}_2, \bar{A}_3$ 中至少有一个发生, 因此可表示为 $\overline{A_1 A_2 A_3}$ 或 $\bar{A}_1 \cup \bar{A}_2 \cup \bar{A}_3$;

(6) “至少中两枪”即 $A_1 A_2, A_1 A_3, A_2 A_3$ 中至少有一个发生, 所以可表示为 $A_1 A_2 \cup A_1 A_3 \cup A_2 A_3$.

§ 3 概率

对于一个随机事件来说, 它在一次试验中可能发生, 也可能不发生. 但人们还是希望知道某些事件在一次试验中发生的可能性究竟有多大. 例如, 在 100 件产品中含有 15 件次品, 从中任取两件, 取到的是合格品的可能性有多大呢? 我们希望找到一个合适的数来表征事件在一次试验中发生的可能性大小. 为此, 我们先引入频率概念, 它描述了事件发生的频繁程度, 进而引出表征事件在一次试验中发生的可能性大小的数——概率.

1.3.1 频率

定义 1 在相同条件下，进行了 n 次试验，在这 n 次试验中，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，并记为 $f_n(A)$ 。

频率具有下述基本性质：

- (1) $0 \leq f_n(A) \leq 1$;
- (2) $f_n(S) = 1$;
- (3) 设 $A_1, A_2, \dots, A_k$ 是两两互不相容的事件，则
$$f_n(A_1 \cup A_2 \cup \dots \cup A_k) = f_n(A_1) + f_n(A_2) + \dots + f_n(A_k).$$

由于事件 A 发生的频率是它发生的次数与试验的次数之比，其大小表示 A 发生的频繁程度。频率大，事件 A 发生就频繁，这意味着事件 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事件 A 在一次试验中发生的可能性的 大小。先看下面的例子。

例 1 考虑“抛硬币”这个试验，我们将一枚硬币抛掷 5 次、50 次、500 次，各做 10 遍，得到数据见表 1-1 和表 1-2。其中 n_H 表示 H 发生的频数， $f_n(H)$ 表示 H 发生的频率。

表 1-1

试验序号	$n=5$		$n=50$		$n=500$	
	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
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	投硬币的次数	出现正面的次数	出现正面的频率
德·摩根	2 048	1 061	0.518 1
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

从投币试验中我们可以总结如下两点:

(1) 频率具有随机波动性, 即使在相同条件下进行 n 次试验, $f_n(A)$ 也会不尽相同;

(2) 抛硬币次数较小时, 频率波动幅度较大. 但是, 随着试验的次数 n 的逐渐增大, 频率 $f_n(A)$ 呈现出稳定性, 逐渐稳定于某个常数.

这种“频率稳定性”即通常所说的统计规律性. 所以我们做大量的重复试验来计算频率 $f_n(A)$, 以它来描述事件 A 发生可能性的大小是合适的.

但是, 在实际中我们不可能对每一个事件都做大量试验, 然后求得事件的频率. 为了理论研究的需要, 我们从频率的稳定性和频率的性质得到启发, 给出如下描述事件发生可能性大小的概率的定义.

1.3.2 概率

事件发生的可能性大小是事件本身所固有的性质. 频率的统计规律性为概率提供了经验基础, 但是不能作为一个严格的数学定义, 1933 年, 苏联著名的数学家柯尔莫哥洛夫在他的《概率论的基本概念》一书中给出了现在已被广泛接受的概率公理化体系, 第一次将概率论建立在严密的逻辑基础上.

定义 2 设 E 是随机试验, S 是它的样本空间, 对于 E 的每一个事件赋予一个实数, 记为 $P(A)$, 称为事件 A 的概率, 这里集合函数 $P(\cdot)$ 满足下列条件:

(1) 非负性 对于每一个事件 A , 有 $P(A) \geq 0$;

(2) 规范性 对于必然事件 S , 有 $P(S) = 1$;

(3) 可列可加性 设 $A_1, A_2, \dots$ 是两两互不相容事件, 即对于 $A_i A_j = \emptyset (i \neq j; i, j = 1, 2, \dots)$, 有 $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$.

由概率的定义, 可以推得概率的一些性质.

性质 1 $P(\emptyset) = 0$.

证明 令 $A_n = \emptyset (n = 1, 2, \dots)$, 则 $\bigcup_{n=1}^{\infty} A_n = \emptyset$, 且 $A_i A_j = \emptyset (i \neq j; i, j = 1, 2, \dots)$. 由概率的可列可加性得

$$P(\emptyset) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(\emptyset).$$

由概率的非负性知, $P(\emptyset) \geq 0$, 故由上式知 $P(\emptyset) = 0$.

性质 2 (有限可加性) 若 $A_1, A_2, \dots, A_n$ 为有限个两两不相容的事件, 则 $P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$.

证明 令 $A_{n+1} = A_{n+2} = \dots = \emptyset$, 即有 $A_i A_j = \emptyset (i \neq j; i, j = 1, 2, \dots)$, 由概率的可列可加性得

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \\ &= \sum_{k=1}^n P(A_k) + 0 \\ &= P(A_1) + P(A_2) + \dots + P(A_n). \end{aligned}$$

性质 3 (减法公式) 设 A, B 是两个事件, 若 $A \subset B$, 则 $P(B - A) = P(B) - P(A)$, 且

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

证明 由 $A \subset B$ 知 $B = A \cup (B - A)$, 且 $A(B - A) = \emptyset$, 再由概率的有限可加性得

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

由概率的非负性可知 $P(B - A) \geq 0$, 故 $P(B) - P(A) = P(B - A) \geq 0$, 从而 $P(B) \geq P(A)$.

一般地, 设 A, B 是任意两个事件, 则 $P(B - A) = P(B) - P(AB)$.

性质 4 对于任意事件 A , 有 $P(A) \leq 1$.

证明 因为 $A \subset S$, 由本节的性质 3 得, $P(A) \leq P(S) = 1$.

性质 5 (逆事件的概率) 对于任意事件 A , 有 $P(\bar{A}) = 1 - P(A)$.

证明 因 $A \cup \bar{A} = S$, 且 $A\bar{A} = \emptyset$, 由有限可加性得 $1 = P(S) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.

性质 6 (加法公式) 设 A, B 是任意两个事件, 则 $P(A \cup B) = P(A) + P(B) - P(AB)$.

证明 因 $A \cup B = A \cup (B - AB)$, 且 $A(B - AB) = \emptyset$, $AB \subset B$, 故由本节的性质 2 和性质 3 得到

$$\begin{aligned} P(A \cup B) &= P(A) + P(B - AB) \\ &= P(A) + P(B) - P(AB). \end{aligned}$$

加法公式还可以推广到多个事件的情况, 例如, 设 A_1, A_2, A_3 是任意三个事件, 则有 $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_2A_3) - P(A_3A_1) + P(A_1A_2A_3)$. 一般地, 对于任意 n 个事件 $A_1, A_2, \dots, A_n$, 可以用数学归纳法证得

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

例 2 已知 $P(A) = 0.4$, $P(B) = 0.3$, $P(A \cup B) = 0.5$, 求 $P(A\bar{B})$.

解 利用和事件概率公式和差事件概率公式, 即可计算.

由于 $P(AB) = P(A) + P(B) - P(A \cup B) = 0.4 + 0.3 - 0.5 = 0.2$,

故有 $P(A\bar{B}) = P(A) - P(AB) = 0.4 - 0.2 = 0.2$.

例 3 某棉麦连作地区, 因受气候条件的影响, 棉花减产的概率为 0.08, 小麦减产的概率为 0.06, 棉花、小麦减产的概率为 0.04, 试求:

(1) 棉花、小麦至少有一样减产的概率;

(2) 棉花、小麦至少有一样不减产的概率.

解 用 A 表示“棉花减产”, B 表示“小麦减产”, 则 $AB = \{\text{棉花、小麦都减产}\}$, $A \cup B = \{\text{棉花、小麦至少有一样减产}\}$, $\bar{A} \cup \bar{B} = \{\text{棉花、小麦至少有一样不减产}\}$,

由 $P(A) = 0.08$, $P(B) = 0.06$, $P(AB) = 0.04$, 得

(1) $P(A \cup B) = P(A) + P(B) - P(AB) = 0.10$;

(2) $P(\bar{A} \cup \bar{B}) = P(\overline{AB}) = 1 - P(AB) = 0.96$.

例 4 假定某校体育活动的统计如下: 有 50% 的学生喜爱田径运动, 50% 的学生喜爱球类运动, 20% 的学生喜爱体操运动, 20% 的学生喜爱田径和球类运动, 10% 的学生喜爱田径和体操运动, 10% 的学生喜爱球类和体操运动, 5% 的学生喜爱这三类运动, 试计算该校不喜爱这三项运动中任何一项的学生所占的百分比.

解 用 A 表示“该学生喜爱田径运动”, B 表示“该学生喜爱球类运动”, 用 C 表示“该学生喜爱体操运动”, D 表示“该学生不喜爱这三项运动”, 则

$$P(A)=50\%, P(B)=50\%, P(C)=20\%, P(AB)=20\%, P(AC)=10\%, \\ P(BC)=10\%, P(ABC)=5\%.$$

根据概率的加法公式, 得

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) = 85\%, \\ P(D) = P(\bar{A} \bar{B} \bar{C}) = P(\overline{A \cup B \cup C}) = 1 - P(A \cup B \cup C) = 15\%.$$

1.3.3 古典概型

古典概型又称为等可能概型, 它在概率论发展初期曾是主要的研究对象, 古典概型的一些概念具有直观、容易理解的特点, 有着广泛的应用.

1. 古典概型的定义

定义 3 若随机试验满足以下两个条件:

- (1) 试验的样本空间只含有有限个样本点;
- (2) 试验中每个样本点发生的可能性相同.

则称这种试验为古典概型.

2. 古典概型的计算方法

下面我们利用概率的性质来讨论等可能概型中事件概率的计算公式.

设试验的样本空间为 $S = \{e_1, e_2, \dots, e_n\}$, e_i 是基本事件, $i=1, 2, \dots, n$. 由于在试验中每个基本事件发生的可能性相同, 即有

$$P(\{e_1\}) = P(\{e_2\}) = \dots = P(\{e_n\}).$$

又由于基本事件是两两互不相容的, 于是

$$1 = P(S) = P(\{e_1\} \cup \{e_2\} \cup \dots \cup \{e_n\}) \\ = P(\{e_1\}) + P(\{e_2\}) + \dots + P(\{e_n\}) = nP(\{e_i\}),$$

故
$$P(\{e_i\}) = \frac{1}{n}, i=1, 2, \dots, n.$$

若事件 A 包含 k 个基本事件, 即 $A = \{e_{i_1}\} \cup \{e_{i_2}\} \cup \dots \cup \{e_{i_k}\}$, 这里 $i_1, i_2, \dots, i_k$ 是 $1, 2, \dots, n$ 中某 k 个不同的数, 则有 $P(A) = \sum_{j=1}^k P(\{e_{i_j}\}) = \frac{k}{n}$. 式中, k 为事件 A 所包含的基本事件数, n 为 S 所包含的基本事件数, 即

$$P(A) = \sum_{j=1}^k P(\{e_{i_j}\}) = \frac{k}{n} = \frac{\text{事件 } A \text{ 包含的基本事件数}}{S \text{ 中包含的基本事件数}}.$$

例 5(产品的随机抽样问题) 已知 12 件产品中有 2 件次品, 从这些产品中任意抽 4 件产品, 求:

- (1) 恰好取得 1 件次品的概率;
- (2) 至少取得 1 件次品的概率.

解 (1) 方法一: 不考虑抽取的次序, 则

S 中含有 $C_{12}^4 = 495$ 个基本事件,

设事件 A 表示“恰取得 1 件次品”, 故 A 包含 $C_{10}^3 C_2^1$ 个基本事件, 从而有

$$P(A) = \frac{C_{10}^3 C_2^1}{C_{12}^4} = \frac{240}{495} \approx 0.485.$$