



“十三五”江苏省高等学校重点教材

概率论 与数理统计

(第二版)

◆◆◆ 张爱武 主编 ◆◆◆



科学出版社

“十三五”江苏省高等学校重点教材

概率论与数理统计

(第二版)

张爱武 主编

江苏高校品牌专业建设工程资助项目 (PPZY2015C211)

盐城师范院校品牌专业建设工程资助项目 (YCTUPPZYC03)

科学出版社

北京

内 容 简 介

本书是“十三五”江苏省高等学校重点教材,内容包括引言、随机事件及概率、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、数理统计的基本概念、参数估计、假设检验、方差分析及线性回归分析初步等内容。书中每节都安排了习题,习题分为两个部分,横线以上部分为基础题,横线以下部分为提高题,以适应不同层次学生的学习需求。每章结尾都增设了本章概要、常用术语和常用公式,帮助学生复习。

本书可作为高等院校理工类及经济类各专业的概率论与数理统计课程教材,也可供教师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 / 张爱武主编. —2 版. —北京: 科学出版社, 2018.8

“十三五”江苏省高等学校重点教材

ISBN 978-7-03-057813-6

I. ①概… II. ①张… 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 (2018) 第 129275 号

责任编辑: 李淑丽 李 萍 / 责任校对: 彭珍珍
责任印制: 霍 兵 / 封面设计: 华路天然工作室

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>

北京市密东印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2018 年 8 月第 一 版 开本: 787×1092 1/16

2018 年 8 月第一次印刷 印张: 20 1/2

字数: 486 000

定价: 57.50 元

(如有印装质量问题, 我社负责调换)

第二版前言

本书的第一版是 2007 年江苏省精品教材立项建设项目的成果, 2011 年被评为江苏省精品教材, 2017 年立项为“十三五”江苏省高等学校重点教材建设项目(修订教材), 2018 年确定为江苏省重点教材, 自 2013 年出版以来, 也得到了广大读者和同行的一致好评, 这使得本书的影响力也越来越大. 面对精品教材和重点教材建设的要求, 面对广大同仁的鼓励, 我们感觉有必要将近几年在教学过程中的尝试进行系统的总结, 于是有了再版的想法.

本次修订主要做了如下修改: ①针对在教学中遇到的问题, 在不影响知识体系的情况下, 我们对部分内容进行了修改, 以便使学生更容易理解; ②为了激发学生学习概率统计的兴趣, 增加了一些实验的经典案例; ③更新了部分例题和习题; ④为了适应教材的需求, 在假设检验、方差分析及回归分析中增加了 Matlab 程序实现; ⑤为了便于学生复习, 在每章后面增加了本章概要、常用术语和常用分式.

本书的分工如下: 第 1、2 章由张爱武执笔, 第 3、4 章由袁兰兰执笔, 第 5、6 章由黄娟娟执笔, 第 7、8 章和实验的经典案例由李万斌执笔, Matlab 程序实现由孙慧慧执笔, 张爱武教授对全书进行了统稿. 在编写过程中, 参阅了不少参考文献, 同时也吸取了不少同行的意见和建议; 这次再版得到了江苏高校品牌专业建设工程资助项目(PPZY2015C211)和盐城师范院校品牌专业建设工程资助项目(YCTUPPZYC03)的资助, 同时得到了科学出版社一如既往的鼓励和支持. 在此一并表示感谢.

本书虽经过多次修改, 限于编者水平, 难免有不足之处, 恳请读者批评指正, 我们将作进一步改进.

编 者

2018 年 4 月

第一版前言

概率论与数理统计是理工类及经济类专业的一门重要的基础课. 为了适应教学的需要, 编者集盐城师范学院概率论与数理统计全体任课教师的多年教学与科研经验, 结合普通高等院校概率论与数理统计课程特点和教学要求编写了本书. 本书是 2007 年江苏省精品教材立项建设项目的成果; 并于 2011 年被评为江苏省精品教材. 概率论是从数量侧面研究随机现象的统计规律性, 它是本课程的理论基础. 本书 1~4 章是概率论的基本内容. 数理统计是处理随机数据, 建立有效的统计方法, 进行统计推断. 本书 5~8 章是数理统计的基本内容.

本书着重介绍了概率论与数理统计的概念、方法, 注意描述概念的实际意义, 并通过精选例题加深对概念的理解, 体现由浅入深、启发诱导的教学方法; 力求在循序渐进的过程中, 使读者逐步掌握概率论与数理统计的基本方法; 突出概率论与数理统计方法的应用, 对较烦琐的理论推导适当降低要求; 本书的习题分节设立, 这样可使习题更具有针对性, 习题数量也明显增加, 在习题的安排上体现层次性, 在每节后面配备的习题分两部分, 一部分为基本要求, 另一部分为提高要求(横线以下部分).

使用本书大约需要 72 学时. 若课时较少, 可选取部分内容组织教学. 譬如, 概率论部分可选取第 1~3 章大部分内容. 统计部分可选取第 5~7 章大部分内容, 其中最小方差无偏估计、两样本的假设检验与区间估计均可略去.

本书由盐城师范学院张爱武编著, 李万斌、孙慧慧、黄娟娟等老师对本书的编写提出了宝贵的意见, 在此对他们深表感谢. 在编写本过程中, 参阅了不少参考文献, 在此对这些文献的作者表示感谢. 本书虽经多次修改, 但限于编者水平, 仍会有缺点和不足. 恳请读者批评指正, 我们将作进一步改进.

编者

2013 年 4 月

目 录

第二版前言	
第一版前言	
引言	1
第 1 章 随机事件及概率	4
1.1 随机事件与样本空间	4
1.1.1 基本事件与样本空间	4
1.1.2 随机事件	5
1.1.3 事件的关系与运算	6
1.1.4 事件域	10
习题 1.1	11
1.2 概率定义及概率的性质	12
1.2.1 概率的描述性定义	12
1.2.2 概率的统计定义	12
1.2.3 概率的公理化定义	14
1.2.4 概率的性质	15
习题 1.2	17
1.3 古典概型与几何概型	18
1.3.1 古典概型	18
1.3.2 几何概型	23
习题 1.3	26
1.4 条件概率的计算公式	27
1.4.1 条件概率	27
1.4.2 乘法公式	28
1.4.3 全概率公式	29
1.4.4 贝叶斯公式	31
习题 1.4	32
1.5 独立性与伯努利概型	33
1.5.1 事件的独立性	33
1.5.2 伯努利概型	37
习题 1.5	39
本章概要	40
常用术语	41
常用公式	41

第 2 章 随机变量及其分布	43
2.1 随机变量及分布函数	43
2.1.1 随机变量及其分类	43
2.1.2 一维随机变量的分布函数	44
2.1.3 多维随机变量的联合分布函数	47
2.1.4 随机变量的独立性	48
习题 2.1	49
2.2 离散型随机变量及其分布列	50
2.2.1 一维离散型随机变量及分布列	50
2.2.2 多维离散型随机变量及其联合分布列	56
2.2.3 离散型随机变量的独立性	59
习题 2.2	61
2.3 连续型随机变量及其分布	63
2.3.1 一维连续型随机变量	63
2.3.2 二维连续型随机变量及其密度函数	69
2.3.3 连续型随机变量的独立性条件	72
习题 2.3	73
2.4 随机变量函数的分布	74
2.4.1 一维随机变量函数的分布	74
2.4.2 连续型随机变量函数的分布	76
习题 2.4	85
2.5 条件分布	87
2.5.1 条件分布的概念	87
2.5.2 离散型随机变量的条件分布	88
2.5.3 连续型随机变量的条件密度	91
习题 2.5	93
本章概要	94
常用术语	94
常用公式	95
第 3 章 随机变量的数字特征	101
3.1 随机变量的数学期望	101
3.1.1 数学期望的概念	101
3.1.2 几种常用分布的期望	105
3.1.3 随机变量函数的数学期望	107
3.1.4 数学期望的性质	110
习题 3.1	113
3.2 随机变量的方差	114
3.2.1 方差的概念	115

3.2.2 几种常用分布的方差	115
3.2.3 方差的性质	118
3.2.4 切比雪夫不等式	120
习题 3.2	121
3.3 协方差、相关系数	122
3.3.1 协方差	122
3.3.2 相关系数	125
3.3.3 矩	129
3.3.4 协方差矩阵	129
3.3.5 n 维正态分布的概率密度	130
习题 3.3	130
3.4 条件期望与条件方差	132
3.4.1 条件期望	132
3.4.2 条件方差	135
习题 3.4	136
本章概要	136
常用术语	137
常用公式	137
第 4 章 大数定律与中心极限定理	140
4.1 大数定律	140
4.1.1 大数定律的意义	140
4.1.2 几种常用大数定律	141
习题 4.1	144
4.2 随机变量序列的两种收敛性	145
4.2.1 依概率收敛	145
4.2.2 依分布收敛	149
习题 4.2	150
4.3 中心极限定理	151
4.3.1 中心极限定理的概念	151
4.3.2 独立同分布的中心极限定理	152
4.3.3 棣莫弗-拉普拉斯中心极限定理	154
习题 4.3	158
本章概要	158
常用术语	159
常用公式	159
第 5 章 数理统计的基本概念	161
5.1 总体与样本	161
5.1.1 总体与个体	161

5.1.2	简单随机样本	162
5.1.3	参数与参数空间	163
习题 5.1		164
5.2	直方图与经验分布函数	164
5.2.1	直方图	164
5.2.2	经验分布函数	165
习题 5.2		167
5.3	统计量及其分布	167
5.3.1	统计量的概念	168
5.3.2	统计量的分布	169
5.3.3	分位数	173
5.3.4	正态总体的抽样分布	174
习题 5.3		176
本章概要		178
常用术语		178
常用公式		178
第 6 章	参数估计	180
6.1	参数的点估计	180
6.1.1	点估计的概念	180
6.1.2	矩法估计	180
6.1.3	极大似然估计	183
习题 6.1		188
6.2	估计量的评价准则	189
6.2.1	无偏性	189
6.2.2	最小方差性和有效性	192
6.2.3	一致性(相合性)	195
习题 6.2		195
6.3	参数的区间估计	197
6.3.1	区间估计的一般步骤	197
6.3.2	单个正态总体参数的区间估计	198
6.3.3	双正态总体参数的区间估计	201
习题 6.3		203
本章概要		204
常用术语		204
常用公式		205
第 7 章	假设检验	206
7.1	假设检验的基本思想和程序	206
7.1.1	假设检验的基本思想	206

7.1.2 假设检验的程序	209
习题 7.1	209
7.2 正态总体参数的假设检验	210
7.2.1 U 检验	211
7.2.2 T 检验	217
7.2.3 χ^2 检验	223
7.2.4 F 检验	226
习题 7.2	228
7.3 检验的实际意义及两类错误	230
7.3.1 检验结果的实际意义	230
7.3.2 检验中的两类错误	231
7.3.3 样本容量确定问题	233
习题 7.3	235
7.4 非参数假设检验	235
7.4.1 χ^2 拟合检验法	236
7.4.2 独立性检验	240
习题 7.4	243
本章概要	244
常用术语	244
常用公式	245
第 8 章 方差分析及线性回归分析初步	246
8.1 方差分析	246
8.1.1 方差分析的基本原理	246
8.1.2 单因子方差分析方法	247
8.1.3 单因子方差分析中的参数估计	254
8.1.4 二因子方差分析	255
习题 8.1	261
8.2 线性回归分析初步	263
8.2.1 回归分析的相关概念	263
8.2.2 一元线性回归	264
8.2.3 多元线性回归	272
习题 8.2	278
本章概要	279
常用术语	280
常用公式	280
部分习题答案	282
参考文献	301
附录 常用分布表	302
附表 1 常用的概率分布、数学期望及方差	302

附表 2	标准正态分布表	303
附表 3	t 分布表	304
附表 4	χ^2 分布表	305
附表 5	F 分布表	307
附表 6	泊松分布表	314
附表 7	相关系数临界值表	315

引 言

一、必然现象与随机现象

在自然界和人的实践活动中经常遇到各种各样的现象,这些现象大体可分为两类:一类是确定的,例如“在一个标准大气压下,纯水加热到 100°C 时必然沸腾”,“向上抛一块石头必然下落”,“同性电荷相斥,异性电荷相吸”等等,这种在一定条件下有确定结果的现象称为必然现象(确定性现象).

另一类现象是随机的,例如,在相同的条件下,向上抛一枚质地均匀的硬币,其结果可能是正面朝上,也可能是反面朝上,不论如何控制抛掷条件,在每次抛掷之前无法肯定抛掷的结果是什么,这个试验多于一种可能结果,但是在试验之前不能肯定试验会出现哪一个结果.同样地,同一门大炮对同一目标进行多次射击(同一型号的炮弹),各次弹着点可能不尽相同,并且每次射击之前无法肯定弹着点的确切位置,以上所举的现象都具有随机性,即在一定条件下进行试验或观察会出现不同的结果(也就是说,多于一种可能的试验结果),而且在每次试验之前都无法预言会出现哪一个结果(不能肯定试验会出现哪一个结果),这种现象称为随机现象.

再看两个试验.

试验 I:一盒中有 10 个完全相同的白球,搅匀后从中摸出一球;

试验 II:一盒中有 10 个相同的球,其中 5 个白球,5 个黑球,搅匀后从中任意摸取一球.

对于试验 I 而言,在球没有取出之前,我们就能确定取出的球必是白球,也就是说在试验之前就能判定它只有一个确定的结果,这种现象就是必然现象(确定性现象).

对于试验 II 来说,在球没有取出之前,不能确定试验的结果(取出的球)是白球还是黑球,也就是说一次试验的结果(取出的球)出现白球还是黑球,在试验之前无法肯定.对于这一类试验而言,骤然一看,似乎没有什么规律而言,但是实践告诉我们,如果从盒子中反复多次取球(每次取一球,记录球的颜色后仍把球放回盒子中搅匀),那么总可以观察到这样的事实,当试验次数 n 相当大时,出现白球的次数 $n_{\text{白}}$ 和出现黑球的次数 $n_{\text{黑}}$ 是很接近的,比值 $\frac{n_{\text{白}}}{n}$ (或 $\frac{n_{\text{黑}}}{n}$) 会逐渐稳定于 $\frac{1}{2}$,出现这个事实是完全可以理解的,因为盒子中的黑球数与白球数相等,从中任意摸一球取得白球或黑球的“机会”相等.

试验 II 所代表的类型,它有多于一种可能的结果,但在试验之前不能确定试验会出现哪一种结果,这类试验所代表的现象称为随机现象,对于试验而言,一次试验看不出什么规律,但是“大数次”地重复这个试验,试验的结果又遵循某些规律,这些规律称为“统计规律”.在客观世界中,随机现象是极为普遍的,例如“某地区的年降雨量”,“某电话交换台在单位时间内收到的用户的呼唤次数”,“一年全省的经济总量”等等,这些都是随机现象.

二、随机试验

如果一个试验满足下述条件:

(1) 试验可以在相同的条件下重复进行(可重复性);

(2) 试验的所有可能结果是明确的、可知道的(在试验之前就可以知道的),并且不止一个(明确性);

(3) 每次试验总是恰好出现这些可能结果中的一个,但在一次试验之前却不能肯定这次试验出现哪一个结果(不确定性).

称这样的试验是一个随机试验,为方便起见,也简称为试验,今后讨论的试验都是指随机试验.

三、概率论与数理统计的研究对象

概率论是从数量侧面研究随机现象及其统计规律性的数学学科,它的理论严谨,应用广泛,并且有独特的概念和方法,同时与其他数学分支有着密切的联系,是近代数学的重要组成部分.

数理统计是对随机现象统计规律归纳的研究,就是利用概率论的结果,深入研究统计资料,观察这些随机现象并发现其内在的规律性,进而作出一定精确程度的判断,将这些研究结果加以归纳整理,形成一定的数学模型.虽然概率论与数理统计在方法上有所不同,但作为一门学科,它们却相互渗透,相互联系.

概率论与数理统计这门学科的应用相当广泛,不仅在天文、气象、水文、地质、物理、化学、生物、医学等学科有其应用,而且在农业、工业、商业、军事、电信部门等也有广泛的应用.

四、概率论与数理统计发展简史

概率论被称为“赌博起家”的理论.

概率论产生于 17 世纪中叶,是一门比较古老的数学学科,有趣的是:尽管任何一门的数学分支的产生与发展都不外乎是生产、科学或数学自身发展的推动,但是概率论的产生,却起始于对赌博的研究,当时两个赌徒约定赌若干局,并且谁先赢 c 局便是赢家,若一个赌徒赢 a 局 ($a < c$),另一赌徒赢 b 局 ($b < c$) 时终止赌博,问应当如何分赌本?最初正是一个赌徒将问题求教于帕斯卡,促使帕斯卡同费马讨论这个问题,从而他们共同建立了概率论的第一基本概念——数学期望.

1657 年惠更斯也给出了一个与他们类似的解法.

在他们之后,对于研究这种随机(或称偶然)现象规律的概率论做出了贡献的是伯努利家族的几位成员,雅各布给出了赌徒输光问题的详尽解法,并证明了被称为“大数定律”的一个定理(伯努利定理),这是研究偶然事件的古典概率论中极其重要的结果,它表明在大量观察中,事件的频率与概率是极其接近的.历史上第一个发表有关概率论论文的人是伯努利,他于 1713 年发表了一篇关于极限定理的论文,概率论产生后的很长一段时间内都是将古典模型作为概率来研究的,直到 1812 年拉普拉斯在他的著作《分析概率论》中给出概率明确的定义,并且还建立了观察误差理论和最小二乘法估计法,从这时开始对概率的研究,实现了从古典概率论向近代概率论的转变.

概率论在 20 世纪再度迅速发展起来,这是由于科学技术发展迫切地需要研究有关一个或多个连续变化着的参变量的随机变数理论,即随机过程论,1906 年俄国数学家马尔可夫(1856~1922)提出了所谓“马尔可夫链”的数学模型,对发展这一理论做出贡献的还有柯尔莫哥洛夫(苏联)、费勒(美国);1934 年苏联数学家辛钦又提出了一种在时间中均匀进行着的

平稳过程的理论. 随机过程理论在科学技术中有着重要的应用, 现在已建立了马尔可夫过程与随机微分方程之间的联系.

1960年, 卡尔曼(1930~2016)建立了数字滤波论, 进一步发展了随机过程在制导系统中的应用. 1917年苏联科学家伯恩斯坦首先给出了概率论的公理体系. 1933年柯尔莫哥洛夫又以更完整的形式提出了概率论的公理结构, 从此, 更现代意义上的完整的概率论臻于完成.

相对于其他许多数学分支而言, 数理统计是一个比较年轻的数学分支, 它是研究怎样去有效地收集、整理和分析带有随机性的数据, 以对所观察的问题作出推断或预测, 直至为采取一定的决策和行动提供依据与建议. 数理统计学是伴随着概率论的发展而发展起来的. 当人们认识到必须把数据看成是来自具有一定概率分布的总体, 所研究的对象是这个总体而不能局限于数据本身之日, 也就是数理统计学诞生之时(确切时间至今难定论).

从现有资料看, 19世纪中叶以前已出现了若干重要工作, 特别是高斯和勒让德关于观测数据误差分析和最小二乘法的研究. 到19世纪末, 经过包括皮尔逊在内的一些学者的努力, 这门学科已经开始形成.

数理统计学发展成一门成熟的学科, 则是20世纪上半叶的事, 它很大程度上要归功于皮尔逊、费希尔等的工作. 特别是费希尔的贡献, 对这门学科的建立起了决定性的作用. 1946年, 克拉默发表的《统计学数学方法》, 是第一部严谨且比较系统的数理统计学著作, 可以把它作为数理统计学进入成熟阶段的标志.

我国的概率论研究起步较晚, 从1957年开始, 先驱者是许宝騄先生. 1957年暑期许老师在北京师范大学举办了一个概率论与数理统计的讲习班, 从此, 我国对概率论与数理统计的研究有了较大的发展. 现在概率论与数理统计是数学系各专业的必修课之一, 也是工科、经济类专业学生的公共课, 许多高校都设了统计学(特别是财经类高校). 近年来, 我国科学家对概率论与数理统计的研究也取得了较大的成果.

第 1 章 随机事件及概率

1.1 随机事件与样本空间

随机事件与样本空间是概率论中的两个最基本的概念.

1.1.1 基本事件与样本空间

对于随机试验来说,我们感兴趣的往往是随机试验的所有可能结果.例如,掷一枚硬币,我们关心的是出现正面还是出现反面这两个可能结果.若观察的是掷两枚硬币的试验,则可能出现的结果有(正、正),(正、反),(反、正),(反、反)四种,如果掷三枚硬币,其结果还要复杂,但还是可以将它们描述出来的,总之为了研究随机试验,必须知道随机试验的所有可能结果.

1. 基本事件

通常,根据我们研究的目的,将随机试验的每一个可能的结果,称为基本事件.因为随机试验的所有可能结果是明确的,从而所有的基本事件也是明确的.例如,在抛掷硬币的试验中“出现反面”“出现正面”是两个基本事件;又如,在掷骰子试验中“出现一点”,“出现两点”,“出现三点”, \dots “出现六点”都是基本事件.

2. 样本空间

基本事件的全体,称为样本空间.也就是试验所有可能结果的全体是样本空间,样本空间通常用大写的希腊字母 Ω 表示, Ω 中的点即基本事件,也称为样本点,常用 ω 表示,有时也用 A, B, C 等表示.

在具体问题中,给定样本空间是研究随机现象的第一步.

例 1.1.1 一盒中有十个完全相同的球,分别有号码 $1, 2, 3, \dots, 10$, 从中任取一球,观察其标号,写出其样本空间.

解 令 $i = \{\text{取得球的标号为 } i\}$, $i = 1, 2, 3, \dots, 10$, 则 $\Omega = \{1, 2, 3, \dots, 10\}$, $\omega_i = \{\text{标号为 } i\}$, $i = 1, 2, \dots, 10$, 其中, $\omega_1, \omega_2, \dots, \omega_{10}$ 为基本事件(样本点).

例 1.1.2 写出下列试验的样本空间:

- (1) 同时抛掷红色与白色的骰子各一粒,记录其向上一面(简称出现)的点数;
- (2) 同时抛掷两粒骰子,记录出现的点数之和.

解 (1) 用有序数组 (i, j) 表示红色骰子出现 i 点,白色骰子出现 j 点,则样本空间可表示为 $\Omega = \{(1, 1), (1, 2), (1, 3), (1, 4), \dots, (6, 5), (6, 6)\}$ 共 36 个样本点;

(2) 试验和(1)类似,但观察的内容不相同,其样本点也不相同,不难看出(2)的样本空间可表示为: $\Omega = \{2, 3, 4, \dots, 12\}$.

例 1.1.1 和例 1.1.2 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 1.1.3 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.这样就可以把样本空间取为 $\Omega = \{0, 1, 2, \dots\}$.

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 1.1.4 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,样本空间是由试验完全确定的.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.

注意 对于一个实际问题或一个随机现象,考虑问题的角度不同,样本空间也可能选择得不同.

例如,掷骰子这个随机试验,若考虑的是出现的点数,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 $\Omega = \{\text{奇数}, \text{偶数}\}$.

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论中值得研究的一个问题.

1.1.2 随机事件

随机试验总有一定的观察目的,除了考察其所有可能结果组成的样本空间外,还需观察其他各种各样的结果.如在例 1.1.1 中,样本空间 $\Omega = \{1, 2, 3, \dots, 10\}$, 研究下面这些问题:

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

其中 A 为一个基本事件,而 B 与 C 则由基本事件所组成.

例如, B 发生(出现)必须而且只需下列样本点 2, 4, 6, 8, 10 之一发生,它由五个基本事件组成.

同样地, C 发生必须而且只需下列样本点 1, 2, 3, 4, 5 之一发生.

A, B, C 这些结果在一次试验中既可能发生,也可能不发生,体现了随机性.这样的结果称为随机事件,简称事件.习惯上用大写英文字母 A, B, C 等表示,在试验中如果出现 A 中包含了某一个基本事件 ω , 则称作 A 发生,并记作 $\omega \in A$.

我们知道,样本空间 Ω 包含了全体基本事件(样本点),而随机事件不过是由某些特征的基本事件组成的,从集合论的角度来看,一个随机事件不过是样本空间 Ω 的一个子集而已.

如例 1.1.1 中 $\Omega = \{1, 2, 3, \dots, 10\}$, 显然 A, B, C 都是 Ω 的子集,它们可以简单地表示为

$$A = \{3\}, \quad B = \{2, 4, 6, 8, 10\}, \quad C = \{1, 2, 3, 4, 5\}.$$

因为 Ω 由所有基本事件组成,所以在一次试验中,必然要出现 Ω 中的某一基本事件,即 $\omega \in \Omega$, 也就是在试验中 Ω 必然要发生(出现),今后用 Ω 表示一个必然事件,它本身就是 Ω 的子集.例如,掷骰子这个随机试验,“掷出的点数不小于 1”,“掷出的点数是自然数”等,它们在每次试验中必然出现,都是必然事件.

相应地,空集 \emptyset 在任意一次试验中不能有 $\omega \in \emptyset$, 也就是说 \emptyset 永远不可能发生,所以称 \emptyset 为不可能事件.又如,在掷骰子这个试验中,“掷出的点数大于 6”,“掷出的点数是负数”等,它们在每次试验中都不会出现,都是不可能事件.

实质上必然事件就是在每次试验中都发生的事件,不可能事件就是在每次试验中都不发生的事件,必然事件与不可能事件的发生与否,已经失去了“不确定性”,即随机性,因而本质上不是随机事件,但为了讨论问题方便,还是将它看作随机事件.

例 1.1.5 一批产品共 10 件,其中 2 件次品,其余为正品,从中任取 3 件,则 $A = \{\text{恰有} 1 \text{ 件正品}\}$,

$B = \{\text{恰有 2 件正品}\}$, $C = \{\text{至少有 2 件正品}\}$, $D = \{\text{3 件中至少有 1 件次品}\}$, 这些都是随机事件, 而 $\Omega = \{\text{3 件中有正品}\}$ 为必然事件, $\emptyset = \{\text{3 件都是正品}\}$ 为不可能事件.

对于这个随机试验来说, 基本事件总数为 C_{10}^3 个.

1.1.3 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再掌握更复杂事件的规律, 为此需要研究事件和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, 且还定义了 Ω 中的一些事件, $A, B, A_i (i = 1, 2, \dots)$ 等, 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完全类似.

1. 事件的包含关系

定义 1.1.1 若事件 A 发生必然导致事件 B 发生, 则称事件 B 包含 A , 或称 A 是 B 的特款, 也称 A 为 B 的子事件, 记作 $A \subset B$ 或 $B \supset A$.

比如前面提到过的 $A = \{\text{球的标号为 6}\}$, 这一事件就导致了事件 $B = \{\text{球的标号为偶数}\}$ 的发生, 因为摸到标号为 6 的球意味着标号为偶数的球出现了, 所以 $A \subset B$.

可以给定义 1.1.1 一个几何解释, 设样本空间 Ω 是一个长方形, 用一个圆表示一个随机事件, 这类图形称为维恩 (Venn) 图. 如图 1-1 所示, A, B 是两个事件, 也就是说, 它们是 Ω 的子集, “ A 发生必然导致 B 发生” 意味着属于 A 的样本点都在 B 中, 由此可见, 事件 $A \subset B$ 的含义与集合论是一致的.

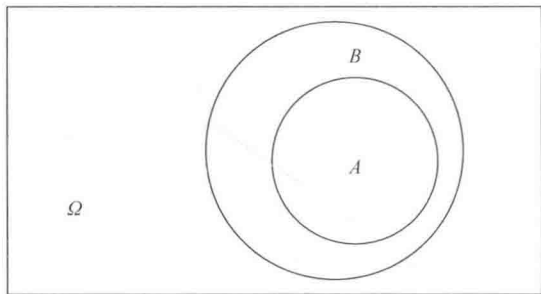


图 1-1

特别地, 对任何事件 A , 有 $A \subset \Omega$, $\emptyset \subset A$.

例 1.1.6 设某种动物从出生活到 20 岁记为 A , 从出生活到 25 岁记为 B , 则 $B \subset A$, 因为某种动物从出生活到 25 岁, 肯定先要活到 20 岁.

2. 事件的相等

定义 1.1.2 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.