

概率论与数理统计

GAILVUNYUSHULITONGJI

■ 高职高专公共课教材

师义民 主编
许勇周 丙 常 副主编



清华大学出版社

高职高专公共课教材

概率论与数理统计

师义民 主 编

许勇周 丙 常 副主编

清华大学出版社

北 京

内 容 简 介

本书共分 8 章, 前 4 章介绍了概率论的基本概念、随机事件及其概率、一维和多维随机变量及其分布、随机变量的数字特征以及大数定律与中心极限定理的内容。第 5~8 章介绍了数理统计学的有关知识, 主要包括数理统计的基本概念、参数估计、假设检验和回归分析等。为了便于阐释和理解概率统计的基本知识, 本书选编了适当难度的例题, 各章均配有习题, 并在书后给出了习题答案。

本书可作为高等职业学校、高等专科学校、成教学院和民办高校学生的使用教材, 也可供工程技术人员参考使用。

本书封面贴有清华大学出版社防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

概率论与数理统计/师义民主编; 许勇周, 丙常副主编. —北京: 清华大学出版社, 2008.7

ISBN 978-7-302-17839-2

I. 概… II. ①师… ②许… ③丙… III. ①概率论—高等学校: 技术学校—教材 ②数理统计—高等学校: 技术学校—教材 IV. TP021

中国版本图书馆 CIP 数据核字(2008)第 084182 号

责任编辑: 刘建龙 桑任松

封面设计: 杨玉兰

版式设计: 北京东方人华科技有限公司

责任印制: 杨 艳

出版发行: 清华大学出版社

地 址: 北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编: 100084

社 总 机: 010-62770175

邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 北京国马印刷厂

经 销: 全国新华书店

开 本: 185×260 印 张: 15.5 字 数: 368 千字

版 次: 2008 年 7 月第 1 版 印 次: 2008 年 7 月第 1 次印刷

印 数: 1~4000

定 价: 23.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题, 请与清华大学出版社出版部联系调换。联系电话: (010)62770177 转 3103 产品编号: 030036-01

前 言

随着科学技术的发展，概率论与数理统计已得到了广泛应用，它已成为高等学校大部分专业必修的一门基础课。本书根据教育部“高职高专教育概率论与数理统计课程教学基本要求”结合作者多年从事本学科教学和科研的体会编写而成，旨在培养学生掌握概率统计的基本思想和方法，提高分析问题和解决问题的能力。

本书以介绍概率论与数理统计的基本知识和方法为主，同时注意它的直观背景和实际意义，力求做到理论与实际相结合，为读者学习后续课程和实际应用，提供了必备的随机数学基础。全书从高职高专教育及工科类教学的特点出发，并考虑到随机数学的特点，力求做到深入浅出，易懂易学。为了便于读者学习，本书以适当的难度循序渐进地选编了一些教学例题，以培养学生对概率统计知识的应用能力。

本书的编写者均为从事概率论与数理统计教学多年的教师。第 1~4 章由师义民编写，第 5~7 章由许勇周编写，第 8 章和各章的习题与答案由丙常编写。李豪亮、胡俊梅和覃晓琼参加了部分工作。全书由师义民统稿、定稿。

限于水平，书中不足之处恳请读者批评指正。

编 者

目 录

第 1 章 随机事件及其概率1	2.4.2 连续型随机变量 函数的分布..... 47
1.1 随机事件.....1	习题 2..... 50
1.1.1 随机现象与随机试验.....1	第 3 章 多维随机变量及其分布 54
1.1.2 随机事件与样本空间.....2	3.1 n 维随机变量及其分类..... 54
1.2 事件间的关系与运算.....2	3.2 二维随机变量及其分布..... 54
1.3 随机事件的概率.....6	3.2.1 联合分布函数..... 54
1.3.1 概率的统计定义.....6	3.2.2 边缘分布函数..... 56
1.3.2 概率的古典定义.....7	3.2.3 二维随机变量的独立性..... 56
1.3.3 概率的性质.....10	3.3 二维离散型随机变量..... 57
1.4 条件概率 全概率公式	3.3.1 联合分布律..... 57
贝叶斯公式.....12	3.3.2 边缘分布律..... 58
1.4.1 条件概率.....12	3.3.3 离散型随机变量的独立性..... 59
1.4.2 全概率公式与贝叶斯公式.....15	3.4 二维连续型随机变量..... 60
1.5 事件的独立性.....17	3.4.1 联合分布密度..... 60
1.5.1 两个事件的独立性.....17	3.4.2 边缘分布密度..... 61
1.5.2 多个事件的独立性.....18	3.4.3 连续型随机变量的独立性..... 62
1.5.3 独立事件概率的计算.....19	3.5 二维随机变量函数的分布..... 65
1.5.4 独立试验序列概型.....20	3.5.1 离散型随机变量函数的 分布..... 65
习题 1.....23	3.5.2 连续型随机变量函数的 分布..... 67
第 2 章 一维随机变量及其分布26	3.5.3 关于极值分布..... 71
2.1 一维随机变量及其分布.....26	习题 3..... 73
2.1.1 随机变量的概念.....26	第 4 章 随机变量的数字特征及 极限定理 76
2.1.2 随机变量的定义.....26	4.1 数学期望..... 76
2.1.3 随机变量的分布函数.....28	4.1.1 数学期望的定义与 计算实例..... 77
2.1.4 分布函数的性质.....28	4.1.2 随机变量函数的数学期望..... 79
2.2 离散型随机变量.....29	4.1.3 数学期望的性质..... 81
2.2.1 分布律与分布函数.....29	4.2 随机变量的方差和矩..... 83
2.2.2 常用的离散型分布.....32	4.2.1 方差的概念..... 83
2.3 连续型随机变量.....36	
2.3.1 分布密度与分布函数.....36	
2.3.2 常用的连续型分布.....39	
2.4 一维随机变量函数的分布.....46	
2.4.1 离散型随机变量 函数的分布.....46	

4.2.2 方差的性质	85	6.5 双正态总体均值差与方差比的	
4.2.3 矩	87	区间估计	133
4.3 协方差与相关系数	88	6.5.1 双正态总体均值差的	
4.3.1 协方差及其性质	88	区间估计	133
4.3.2 相关系数及其性质	90	6.5.2 双正态总体方差比的	
4.4 极限定理概述	93	区间估计	135
4.4.1 切比雪夫不等式	93	6.5.3 单侧置信区间	137
4.4.2 大数定律	94	习题 6	139
4.4.3 中心极限定理	97	第 7 章 假设检验	142
习题 4	99	7.1 假设检验的基本概念	142
第 5 章 数理统计的基本概念与		7.1.1 假设检验的基本原理	143
抽样分布	103	7.1.2 拒绝域与临界值	144
5.1 基本概念	103	7.1.3 假设检验的两类错误与	
5.1.1 总体与样本	103	基本步骤	145
5.1.2 样本的分布	104	7.1.4 单侧检验与双侧检验	146
5.1.3 统计量	105	7.2 单正态总体参数的假设检验	146
5.2 抽样分布	108	7.2.1 正态总体均值的假设检验	146
5.2.1 χ^2 分布	108	7.2.2 均值未知时正态总体方差的	
5.2.2 t 分布	110	检验(χ^2 检验)	148
5.2.3 F 分布	110	7.3 双正态总体参数的假设检验	149
5.2.4 概率分布的分位数	112	7.3.1 双正态总体均值的检验	150
5.3 正态总体的抽样分布	113	7.3.2 双正态总体方差的检验	
习题 5	116	(F 检验)	152
第 6 章 参数估计	119	7.3.3 单侧检验	155
6.1 矩估计	119	7.4 总体分布的假设检验	157
6.2 最大似然估计	121	习题 7	160
6.2.1 似然函数	122	第 8 章 回归分析	163
6.2.2 最大似然估计法	122	8.1 一元线性回归分析	163
6.3 优良性准则	125	8.1.1 引例	163
6.3.1 无偏性	125	8.1.2 一元线性回归模型	164
6.3.2 有效性	126	8.1.3 确定回归直线方程	165
6.3.3 相合性(或一致性)	127	8.1.4 估计量的性质和分布	167
6.4 单正态总体参数的区间估计	128	8.1.5 回归方程显著性检验	169
6.4.1 正态分布总体均值的		8.1.6 预测和控制	172
置信区间	129	8.2 可线性化的非线性回归模型	175
6.4.2 正态总体方差的区间估计	132	8.3 多元线性回归分析	178
		8.3.1 数学模型	178

8.3.2 未知参数的估计	179	附录	194
8.3.3 估计量的分布及性质.....	186	参考答案	223
8.3.4 回归系数和回归方程的 显著性检验	188	参考文献	237
习题 8	190		

第 1 章 随机事件及其概率

1.1 随机事件

1.1.1 随机现象与随机试验

在自然界和人类活动中普遍存在着两类现象. 一类是在一定条件下必然发生的现象, 称这类现象为必然现象或确定性现象. 譬如:

例 1.1 异性电荷相互吸引, 同性电荷相互排斥.

例 1.2 向空中抛掷一个小球, 必然往下落.

另一类现象则是在一定条件下可能发生也可能不发生的现象, 具有不确定性. 对这种现象人们事先无法预知其结果. 譬如:

例 1.3 某篮球运动员投篮一次, 其结果可能命中也可能不中.

例 1.4 上抛一枚均匀的硬币, 落下后可能正面向上, 也可能反面向上.

例 1.5 一批新产品投放市场, 可能畅销也可能滞销.

例 1.6 在某公交车站上, 某一固定的时刻候车的人数.

对于随机现象, 在少数几次试验或观察中其结果无规律性, 但通过长期观察或大量的重复试验可以看出, 试验的结果呈现出一种规律性, 这种规律性称为统计规律性, 它是随机现象自身所具有的特征. 概率论是研究随机现象及其统计规律性的一门数学学科, 它被广泛应用于自然科学、社会科学的许多领域.

为了深入研究随机现象, 就必须在一定的条件下对它进行多次观察. 若把一次观察视为一次试验, 观测到的结果就是试验结果. 概率论中把满足下列条件的试验称为随机试验.

- (1) 试验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, 但试验的所有可能结果事先是明确的.
- (3) 在试验结果揭晓前, 无法预言会发生哪一个可能结果.

本书以后所指的试验若无特别声明, 均指随机试验. 例如:

- (1) 在一定的条件下进行射击练习, 考虑中靶的环数.
- (2) 掷一颗均匀的骰子, 考虑出现的点数.
- (3) 记录某汽车站某时段内候车的人数.

随机现象的表现, 即随机试验的结果数学模型化, 可用集合的概念描述. 例如, 在掷硬币试验中, 试验的结果有两个: “正面”和“反面”如果用 ω_1 表示结果为正面, ω_2 表示结果为反面, 则可以用 $A_1 = \{\omega_1\}$, $A_2 = \{\omega_2\}$ 来表示试验结果. 再考虑射击练习的例子,

若以 ω_i 表示中靶的环数为 i , $i=1,2,\dots,10$. 则 $A_i = \{\omega_i\}$ 可以表示中靶的环数为 i 的结果. $A = \{\omega_1, \omega_3, \omega_5, \omega_7, \omega_9\}$, $B = \{\omega_2, \omega_4, \omega_6, \omega_8, \omega_{10}\}$ 分别表示中靶的环数为奇数和偶数.

1.1.2 随机事件与样本空间

在概率论中, 随机事件的数学描述由以下定义给出.

定义 1.1 对于随机试验, 把每一个可能出现的结果称为样本点, 把某些样本点构成的集合称为随机事件, 简称事件. 把单个样本点构成的集合称为基本事件. 把所有样本点构成的集合称为必然事件或样本空间, 记为 Ω .

通常用 A, B, C, \dots 或 A_1, A_2, A_3, \dots 等大写字母表示随机事件. 为了以后运算封闭, 规定不含任何元素的空集也为事件, 称为不可能事件, 用 \emptyset 表示.

例 1.7 已知一批产品共 30 件, 其中有正品 26 件、次品 4 件, 从中随机取出 5 件. 则

$A_i =$ “被取出的 5 件产品中恰好有 i 件次品”, $i=0,1,2,3,4$;

$B =$ “被取出的 5 件产品中最多有 3 件次品”;

$C =$ “被取出的 5 件产品中正品不超过 2 件”.

这些都是随机事件.

例 1.8 写出掷骰子试验的样本点、样本空间、基本事件、事件 A ——出现偶数、事件 B ——出现奇数.

解: 用 ω_i 表示出现的点数为 i , $i=1,2,\dots,6$; 则 $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$; 基本事件 $A_i = \{\omega_i\}$; 事件 $A = \{\omega_2, \omega_4, \omega_6\}$, $B = \{\omega_1, \omega_3, \omega_5\}$.

例 1.9 从一批灯泡中任意抽取一个, 测试其寿命(用 t 表示, $0 \leq t < +\infty$), 其样本点有无穷多个(且不可数), 样本空间可简记为 $\Omega = \{t | 0 \leq t < +\infty\} = [0, +\infty)$.

1.2 事件间的关系与运算

在概率论中, 人们往往不仅要研究随机试验的一个事件, 还要研究多个事件, 而这些事件之间又有一定的联系. 为了表述事件间的联系, 下面定义事件间的关系和运算.

1. 包含与相等

若事件 A 发生必导致事件 B 发生, 则称事件 B 包含事件 A , 记为 $A \subset B$ 或 $B \supset A$.

这里事件 A “发生”一词是指事件 A 所含的任一样本点出现. 例如掷骰子, 称事件 A ——出现偶数点发生, 指在一次观察(一次投掷)中出现的点数为 2 或 4 或 6.

若事件 B 包含事件 A , 同时事件 A 也包含事件 B , 即 $A \subset B$ 和 $A \supset B$ 同时成立, 则称

事件 A 与 B 相等, 记为 $A=B$.

例如, 从一副 52 张的扑克牌中任取 4 张, 令事件 A 表示“取到至少有 3 张红桃”的事件; B 表示“取到至多有 1 张不是红桃”的事件. 显然 $A=B$. 两事件相等意味着它们由相同的样本点构成, 在试验中必将同时发生. 对于任一事件 A , 有 $\emptyset \subset A \subset \Omega$. 其中 $\emptyset \subset A$ 作为规定而引入.

2. 事件的和 (或并)

“事件 A 与 B 至少发生一个”也是一个事件, 称为 A 与 B 的和(或并), 记为 $A \cup B$.

和事件的概念可推广到有限个或可列多个事件的情形. 一般地“ A_1, A_2, \dots, A_n 中至少发生一个”也是一个事件, 称为 A_1, A_2, \dots, A_n 的和(或并)事件, 记作 $A_1 \cup A_2 \cup \dots \cup A_n$, 或简记为 $\bigcup_{i=1}^n A_i$; 而“可列多个事件 $A_1, A_2, \dots, A_n, \dots$ 至少发生一个”也是事件, 称为可列多个事件 $A_1, A_2, \dots, A_n, \dots$ 的和(或并)事件, 记作 $\bigcup_{i=1}^{\infty} A_i$. 因此, 两事件 A 与 B 的和是指“事件 A 发生, 或 B 发生, 或 A 与 B 都发生”.

在随机事件的讨论中, 应特别注意一些关键词语, 如“或者”、“同时”、“至少”, “至多”等, 它们表述了不同的运算. 在随机事件的讨论中, 应特别注意一些关键词语, 如“或者”、“同时”、“至少”、“至多”等, 它们表述了不同的运算.

3. 事件的积(或交)

“两事件 A 与 B 同时发生”是一个事件, 称为事件 A 与 B 的积(或交)事件, 记为 $A \cap B$ 或 AB . 类似事件的和、积事件的概念可推广到有限个或可列多个事件的情形, 即若 n 个事件“ A_1, A_2, \dots, A_n 同时发生”, 称为 n 个事件 A_1, A_2, \dots, A_n 的积(或交)事件, 记作 $A_1 \cap A_2 \cap \dots \cap A_n$, 简记作 $\bigcap_{i=1}^n A_i$; 可列多个事件“ $A_1, A_2, \dots, A_n, \dots$ 同时发生”也是一个事件, 称为可列多个事件 $A_1, A_2, \dots, A_n, \dots$ 的积(或交)事件, 简记作 $\bigcap_{i=1}^{\infty} A_i$.

4. 事件的互斥(互不相容)

若两事件 A 与 B 不能同时发生, 即满足 $A \cap B = \emptyset$, 则称 A 与 B 是互斥(或互不相容)事件. 若两事件不互斥, 则称为相容. 事件 A 与 B 互斥意味着它们没有公共的样本点. 当事件 A 与 B 互不相容时, 可将它们的和 $A \cup B$ 记为“直和” $A+B$.

对于 n 个事件 A_1, A_2, \dots, A_n , 若其中任意两个事件不能同时发生, 即有

$$A_i A_j = \emptyset, \quad i \neq j, i, j = 1, 2, \dots, n$$

成立, 则称 A_1, A_2, \dots, A_n 两两互斥. 对于无限可列多个事件, 也可仿此建立两两互斥的概念.

5. 事件的差与逆(对立)

“事件 A 发生而 B 不发生”是一个事件，称为事件 A 与 B 的差，记为 $A-B$ 。若事件 A 与 B 满足 $AB=\emptyset$ ，且 $A\cup B=\Omega$ ，则称事件 A 与 B 是对立(互逆)事件。事件 A 与 B 对立意味着它们无公有样本点，而所有样本点又恰好充满样本空间 Ω 。 A 的对立事件记为 \bar{A} 。两事件对立一定互斥，反之则不一定成立。

为了便于理解，人们把事件间的关系和运算用图形示意出来(见图 1.1)，这种图称为文氏图(Venn)。

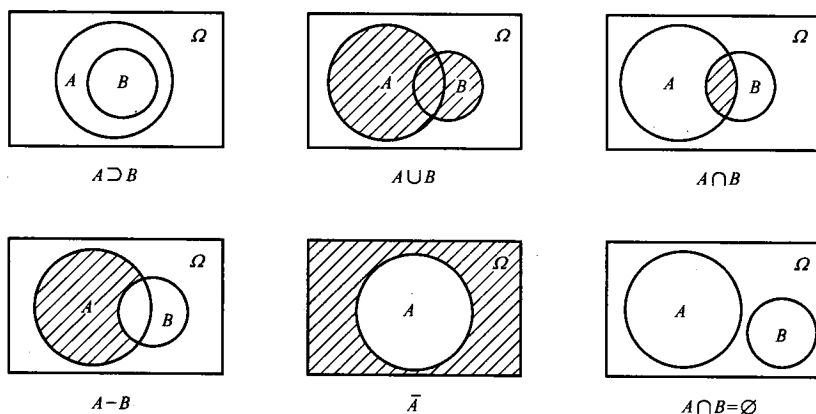


图 1.1 文氏图(Venn)

由上述定义可以推出，事件间具有以下基本关系：

- (1) $\bar{\bar{A}}=A$, $A\cup\bar{A}=\Omega$;
- (2) 若 $A\supset B$, 则 $A\cup B=A$, $AB=B$;
- (3) $\overline{A-B}=\bar{A}\cup B$, $A-B=A\bar{B}=A-AB$;
- (4) $A\cup B=A\cup\bar{A}B=A\cup(B-A)$;
- (5) $AB\subset A\subset A\cup B$, $AB\subset B\subset A\cup B$.

6. 完备事件组

若 n 个事件 A_1, A_2, \dots, A_n 满足 $A_i A_j = \emptyset$, $i \neq j, i, j = 1, 2, \dots, n$, 且 $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$, 则称事件 A_1, A_2, \dots, A_n 构成了一个完备事件组, 或称 A_1, A_2, \dots, A_n 是样本空间 Ω 的一个划分. 显然, A 和 \bar{A} 构成了一个完备事件组.

由集合的运算规律, 可以给出事件间的运算规律. 设 A, B, C 是同一随机试验 E 中的三个随机事件. 则有以下定律:

- (1) 交换律: $A\cup B=B\cup A$, $A\cap B=B\cap A$.
- (2) 结合律: $(A\cup B)\cup C=A\cup(B\cup C)$, $(A\cap B)\cap C=A\cap(B\cap C)$.

(3) 分配律: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(4) 对偶律: $\overline{A \cup B} = \bar{A} \bar{B}$; $\overline{AB} = \bar{A} \cup \bar{B}$, $\overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \bar{A}_i$, $\overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i$.

对于上述运算律的证明,需用严格的集合论证明手法.

由于事件是集合,因而事件的运算即是集合运算.表 1.1 给出了概率论中事件及其运算与集合论中集合及其运算的术语.

表 1.1 术语对照表

记号	概率论	集合论
Ω	样本空间, 必然事件	空间(全集)
\emptyset	不可能事件	空集
ω	样本点	元素
A	随机事件	子集
\bar{A}	事件 A 的对立事件	A 的补集
$A \subset B$	事件 A 发生必然导致事件 B 发生	A 是 B 的子集
$A = B$	事件 A 与事件 B 相等	A 集合与 B 集合相等
$A \cup B$	事件 A 与事件 B 的和	A 集合与 B 集合的并集
$A \cap B$	事件 A 与 B 的积事件	A 集合与 B 集合的交集
$A - B$	事件 A 与事件 B 的差	A 与 B 两集合的差集
$A \cap B = \emptyset$	事件 A 与 B 事件互不相容	A 与 B 两集合中没有相同的元素

例 1.10 设 A, B, C 为三个事件, 则:

(1) A, B, C 都发生可表示为 ABC .

(2) A, B 发生而 C 不发生可表示为 $ABC\bar{C} = AB - C$.

(3) A, B, C 都不发生可表示为 $\bar{A}\bar{B}\bar{C}$.

(4) A, B 中至少有一个发生而 C 不发生可表示为 $(A \cup B)\bar{C} = A \cup B - C$.

(5) A, B, C 中至少有一个发生可表示为 $A \cup B \cup C$.

或 $A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C + A\bar{B}C + \bar{A}BC + \bar{A}B\bar{C} + ABC$.

(6) 事件 A, B, C 中至多有一个发生可表示为 $\bar{A}\bar{B}C \cup \bar{A}B\bar{C} \cup A\bar{B}\bar{C} \cup \bar{A}\bar{B}\bar{C}$.

例 1.11 将两颗均匀的骰子各掷一次, 若以 (x, y) 表示其结果, 其中 x 表示第一颗骰子出现的点数, y 表示第二颗骰子出现的点数, 则样本空间为

$$\Omega = \{(x, y) : x, y = 1, 2, 3, 4, 5, 6\}$$

若以 A, B, C, D 分别表示事件“点数之和等于 2”、“点数之和等于 5”、“点数之和超过 9”、“点数之和不小于 4 也不超过 6”. 则

$$A = \{(1, 1)\}, \quad B = \{(1, 4), (2, 3), (3, 2), (4, 1)\}, \quad C = \{(4, 6), (5, 5), (6, 4), (5, 6), (6, 5), (6, 6)\};$$

$$D = \{(1, 3), (2, 3), (3, 1), (1, 4), (2, 3), (3, 2), (4, 1), (1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}.$$

1.3 随机事件的概率

随机事件在一次试验中可能发生,也可能不发生,其结果往往事先无法预言.但是,在做大量重复试验时,随机事件发生的可能性大小是客观存在的,是事件本身的固有属性.对此,人们试图用一个数加以度量.通俗地说,把度量事件 A 在试验中发生的可能性大小的数叫做概率,并记为 $P(A)$.这样,较大的概率 $P(A)$ 预示着相应事件 A 发生的可能性较大;反之,则可能性较小.

对于给定的随机事件,如何定义和获得该事件发生的概率,通常与试验条件有关.下面只给出在实际中用得较多的两个定义;概率的统计定义和古典概型定义.至于几何定义与公理化定义本书不再涉及.

1.3.1 概率的统计定义

统计定义是以大量重复试验为前提的.为此,首先引入频率的概念,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——概率及其稳定性概念.

定义 1.2 在 N 次重复试验中,事件 A 发生的次数 M (频数)与试验次数 N 的比 $\frac{M}{N}$,称为事件 A 的频率,并记为 $f_N(A)$,即 $f_N(A) = \frac{M}{N}$.

容易验证,频率 $f_N(A)$ 具有如下性质:

- (1) 对于任一事件 A , 有 $0 \leq f_N(A) \leq 1$.
- (2) $f_N(\Omega) = 1$, $f_N(\emptyset) = 0$.
- (3) 若 A_1, A_2, \dots, A_n 是两两互不相容的事件, 则

$$f_N(A_1 \cup A_2 \cup \dots \cup A_n) = f_N(A_1) + f_N(A_2) + \dots + f_N(A_n)$$

实践表明,当试验次数 N 逐渐增大时,频率 $f_N(A)$ 虽然不尽相同,但却稳定在某个实数附近,这种性质称为频率的稳定性.表 1.2 中给出了掷硬币试验的结果,其中 N 表示重复投掷次数, M 表示“正面朝上”(设为事件 A)出现的次数, $f_N(A)$ 表示 N 次试验中事件 A 出现的频率.

$f_N(A)$ 呈现出来的稳定趋势充分说明频率具有稳定性.

频率的稳定性为概率概念的量化提供了可行性途径.所谓统计概型就是以大量重复试验为前提,以频率的稳定性事实为依据给概率以数量刻画的随机试验模型.

定义 1.3 (概率的统计定义) 在相同条件下重复进行 N 次随机试验,如果事件 A 的频率 $f_N(A)$ 随着试验次数 N 的增加而稳定地在某个常数 $p(0 \leq p \leq 1)$ 附近摆动,则称 p 为事件 A 的概率,记为 $P(A)$.

由统计定义求得的概率简称为统计概率.

定义中的大量试验往往难以办到, 实际操作时只是根据需求和可能, 用一定数量试验下的频率作为频率的近似. 例如, 在表 1.2 中给出的硬币投掷试验, 由于当 $N=500$ 时, 频率出现了稳定的趋势, 于是便可认定 $P(A)=0.5$.

表 1.2 均匀硬币在历次投掷中的频率表

试验号	N=5		N=50		N=500	
	M	$f_N(A)$	M	$f_N(A)$	M	$f_N(A)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.5	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1	25	0.5	253	0.506
5	2	0.4	21	0.42	246	0.492
6	4	0.8	18	0.36	244	0.488

在统计概型下, 概率是作为频率的稳定值而引入的, 因而, 概率也应当具备频率的基本性质.

性质 1.1 (概率统计定义的性质)

- (1) 对于任意事件 A , $0 \leq P(A) \leq 1$.
- (2) $P(\Omega)=1$; $P(\emptyset)=0$.
- (3) 若 A_1, A_2, \dots, A_n 是两两互不相容的事件, 则

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

1.3.2 概率的古典定义

概率的统计定义直观地描述了事件发生的可能性大小, 反映了概率的本质内容, 但也有明显的不足, 即在统计定义下求概率要求试验量很大, 这在实际中往往很难办到, 为此引入另一个常用概型——古典概型.

古典概型是古典概率模型的简称. 一般称具有下列两个特征的随机试验模型为古典概型.

- (1) 随机试验只有有限个可能的结果, 即有限个样本点(有限性).
- (2) 每一个样本点发生的可能性相等(等可能性).

古典概型又称为等可能性概型. 在概率论产生和发展的过程中, 它是最早的研究对象, 在实际应用中它也是最常用的一种概率模型.

对于古典概型, 以 $\Omega = \{\omega_1, \dots, \omega_n\}$ 表示样本空间, $\omega_i (i=1, 2, \dots, n)$ 表示样本点, 对于任

一随机事件 $A = \{\omega_1, \dots, \omega_m\}$, 下面给出古典概型的定义.

定义 1.4 (概率的古典概型定义)

对于给定的古典概型, 若样本空间中有 n 个样本点, 事件 A 含有 m 个样本点, 则事件 A 的概率为

$$P(A) = \frac{m}{n} = \frac{\text{事件}A\text{所含样本点的个数}}{\text{样本空间}\Omega\text{所含样本点的个数}} \quad (1.1)$$

显然, 凡属古典概型的问题只需运用式(1.1), 并依据乘法原理或加法原理的思路, 借助排列组合为工具, 直接计算其概率. 由古典定义求得的概率简称为古典概率.

性质 1.2 (古典概率的性质)

- (1) 对于任意事件 A , $0 \leq P(A) \leq 1$.
- (2) $P(\Omega) = 1$, $P(\emptyset) = 0$.
- (3) 若 A_1, A_2, \dots, A_n 是两两互不相容的事件, 则

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

例 1.12 某种产品共有 30 件, 其中含正品 23 件, 次品 7 件, 从中任取 5 件. 试求被取出的 5 件中恰好有 2 件是次品的概率.

解: 设 $A =$ “被取出的 5 件中恰好有 2 件是次品”. 由题设 “从中任取 5 件” 应理解为 “一次取出 5 件”, 故样本点总数 $n = C_{30}^5$. 事件 A 包含的样本点数 $m = C_7^2 C_{23}^3$, 则所求概率为

$$P(A) = \frac{C_7^2 C_{23}^3}{C_{30}^5} = 0.2610.$$

本例的一般情形为: 某件产品共 N 件, 其中含次品 M 件, 从中任取 n 件, 其中恰有 m 件次品(设为事件 A) 的概率为

$$P(A) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \quad (1.2)$$

式中 $m = 0, 2, \dots, \min\{M, n\}$. 在概率论中式(1.2)被称为超几何公式.

例 1.13 某口袋中有 6 只球, 其中 4 只白球, 两只红球, 从袋中取球两次, 每次随机地取一只. 考虑两种取球方式.

① 第一次取一只球, 观察其颜色后放回袋中, 搅匀后再取一球. 这种取球方式叫做有放回取球.

② 第一次取一只球不放入袋中, 第二次从剩余的球中再取一只球. 这种取球方式叫做无放回取球.

试分别就上面两种情况求:

- (1) 取到的两只球都是白球的概率;
- (2) 取到的两只球颜色相同的概率.

解: (1) 令 A_1 表示事件“取到的两只球都是白球”，则

$$\text{有放回取球: } P(A_1) = \frac{4}{6} \times \frac{4}{6} = \frac{4}{9}$$

$$\text{无放回取球: } P(A_1) = \frac{4}{6} \times \frac{3}{5} = \frac{2}{5}$$

(2) 令 A_2 表示事件“取到的两只球颜色相同”，则

$$\text{有放回取球: } P(A_2) = \frac{4}{6} \times \frac{4}{6} + \frac{2}{6} \times \frac{2}{6} = \frac{5}{9}$$

$$\text{无放回取球: } P(A_2) = \frac{4}{6} \times \frac{3}{5} + \frac{2}{6} \times \frac{1}{5} = \frac{7}{15}$$

例 1.14 将含有 3 名优秀生的 15 名学生平均地分成 5 人小组，求下列事件的概率：

(1) A ：每个小组各有一名优秀生。

(2) B ：3 名优秀生被分到一个小组。

(3) C ：3 名优秀生中有 2 名被分到一个小组，另一名分到另一个小组。

解: 15 名学生随机地分成 3 个 5 人小组所有不同的分法为 $\frac{15!}{5!5!5!}$ ，这是基本事件总数。

这些基本事件的发生是等可能的。

(1) A 的有利场合数：将 3 名优秀生分到 3 个小组，每人一组的分法总数为 $3!$ ；12 名其他学生分成 3 组，4 人一组的的所有分法为 $\frac{12!}{4!4!4!}$ ，故 A 的有利场合数为 $3! \times \frac{12!}{4!4!4!}$ 。

于是，所求概率为 $P(A) = \frac{3! \times 12!}{4!4!4!} \bigg/ \frac{15!}{5!5!5!} = \frac{25}{91}$ 。

(2) B 的有利场合数：将 12 名其他学生分成 2 人、5 人、5 人的 3 个小组的所有分法为 $\frac{12!}{2!5!5!} + \frac{12!}{5!2!5!} + \frac{12!}{5!5!2!}$ ，哪一组为 2 人，3 名优秀生就分到这个组，只有一种分法。故

B 的有利场合数为 $3 \times \frac{12!}{5!5!2!}$ 。于是所求概率为 $P(B) = \frac{12! \times 3}{2!5!5!} \bigg/ \frac{15!}{5!5!5!} = \frac{6}{91}$ 。

(3) C 的有利场合数：将 12 名其他学生分成 3 人、4 人、5 人的 3 个小组的所有分法为 $\frac{12!}{3!4!5!} \times 3$ ，将 3 名优秀生分成 2 人、1 人的 2 个组有 C_3^2 种分法，对于每一种这样的分

法将 2 名优秀生分到 3 人组的那个组里，将另一名优秀生分到 4 人组的那个组里，只有一种分法。故 C 的有利场合数为 $\frac{12!}{3!4!5!} \times 3! \times C_3^2$ 。故所求概率为 $P(C) = \frac{12! \times 3! \times C_3^2}{3!4!5!} \bigg/ \frac{15!}{5!5!5!} = \frac{60}{91}$ 。

例 1.15 有 n 个人，每个人都以同样的概率 $\frac{1}{N}$ 被分配在 N ($n \leq N$) 间房中的每一间

中，试求下列各事件的概率：

(1) 某指定的 n 间房中各有一人。

(2) 恰有 n 间房，其中各有一人。

(3) 某指定的一间房中恰有 m ($m \leq n$) 人。

解: 先求样本空间中所含样本点的个数.

首先, 把 n 个人分到 N 间房中去共有 N^n 种分法; 其次, 求每种情形下事件所含的样本点个数.

- (1) 某指定的 n 间房中各有一人, 所含样本点的个数, 即可能的分法为 $n!$.
- (2) 恰有 n 间房中各有一人, 所有可能的分法为 $C_N^n n!$.
- (3) 某指定的一间房中恰有 m 人, 可能的分法为 $C_n^m (N-1)^{n-m}$.

于是可以得到三种情形下事件的概率分别如下:

- (1) $\frac{n!}{N^n}$.
- (2) $\frac{C_N^n \cdot n!}{N^n}$.
- (3) $\frac{C_n^m (N-1)^{n-m}}{N^n}$.

在上述分房问题中, 若令 $N=365$, $n=30$, $m=2$ 则可演化为生日问题.

全班有学生 30 人, 求下列事件的概率:

- (1) 某月指定为 30 天, 每位学生生日各占一天.
- (2) 全班学生生日各不相同.
- (3) 全年某天, 恰有两人在这一天同生日.

利用上述结论可得到概率分别如下:

- (1) $\frac{30!}{365^{30}}$.
- (2) $\frac{C_{365}^{30} \cdot 30!}{365^{30}} \approx 0.294$.
- (3) $\frac{C_{30}^2 (364)^{28}}{(365)^{30}}$.

1.3.3 概率的性质

性质 1.3 根据随机事件概率的定义, 可得到随机事件的概率具有以下性质:

- (1) $P(\emptyset) = 0$.
- (2) 若 A_1, A_2, \dots, A_n 是两两互不相容的事件, 则

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

- (3) $P(\bar{A}) = 1 - P(A)$.
- (4) 若 $B \subset A$, 则 $P(A-B) = P(A) - P(B)$, 且 $P(B) \leq P(A)$.
- (5) 对于任意事件 A, B , $P(A \cup B) = P(A) + P(B) - P(AB)$.

证明: (1) $\Omega = \Omega + \emptyset + \emptyset + \emptyset + \dots$