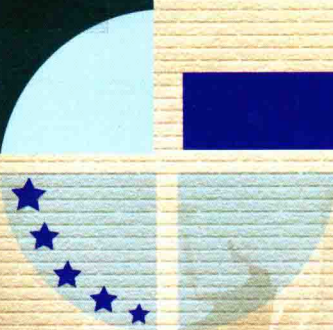




普通高等教育“十三五”规划教材

概率统计 应用与实验

主 编 黄龙生 黄 敏
副主编 宋红凤 许芳忠 夏慧珠



中国水利水电出版社
www.waterpub.com.cn

普通高等教育“十三五”规划教材

概率统计 应用与实验

主 编 黄龙生 黄 敏
副主编 宋红凤 许芳忠 夏慧珠



中国水利水电出版社
www.waterpub.com.cn

· 北京 ·

内 容 提 要

本书内容包括随机事件及其概率、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、常用分布、大数定律及中心极限定理、数理统计基础、参数估计、假设检验、方差分析和回归分析。每章后面有选择题和填空题,适量的应用计算题,且在数理统计部分(第六~第十章)每章配有案例分析,基于 EXCEL 环境设置了常用随机变量分布、参数的区间估计、假设检验、方差分析和回归分析的实验。

本书可作为普通高等学校本科非数学类应用型专业的“概率论与数理统计”课程的教材或参考书,也可用作专科或高职院校相关专业的“概率论与数理统计”课程的教材或参考书,同时也可作为工程技术人员和科技工作者的参考书。

图书在版编目(CIP)数据

概率统计应用与实验 / 黄龙生, 黄敏主编. — 北京:
中国水利水电出版社, 2018.1
普通高等教育“十三五”规划教材
ISBN 978-7-5170-6189-2

I. ①概… II. ①黄… ②黄… III. ①概率统计—高
等学校—教材 IV. ①O211

中国版本图书馆CIP数据核字(2017)第326212号

书 名	普通高等教育“十三五”规划教材 概率统计应用与实验 GAILÜ TONGJI YINGYONG YU SHIYAN
作 者	主 编 黄龙生 黄 敏 副主编 宋红凤 许芳忠 夏慧珠
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: sales@waterpub.com.cn 电话: (010) 68367658 (营销中心)
经 售	北京科水图书销售中心(零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	中国水利水电出版社微机排版中心
印 刷	天津嘉恒印务有限公司
规 格	184mm×260mm 16开本 20印张 474千字
版 次	2018年1月第1版 2018年1月第1次印刷
印 数	0001—2500册
定 价	56.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

前言

“概率论与数理统计”是研究随机现象及其统计规律性的一门学科，在各行各业中有大量的实际应用。普通高校中很多应用型专业，要求学生在掌握概率论与数理统计的基本理论、基本方法和基本技能的基础上，还应加强概率论与数理统计的应用技能，为了适应这种需要，我们编写了这本《概率统计应用与实验》教材。本书强化了概率论与数理统计的应用，以求随机事件的概率及随机变量的数字特征为主线编排概率论部分的内容；在数理统计部分，以扩展内容的形式，增加了部分实用的统计工具与方法的介绍。本书可作为高校非数学类专业的教材或参考书，也可供相关科技工作者参考。

全书力求突出概率论与数理统计的基本思想和方法。本书第一部分用“引例（日常生产生活中的问题）”的方式导入新的概念、思想和方法，力求通俗易懂；对专业术语给出了相应的英语译文，为学生阅读外文资料提供便利；在例题和习题的选取上注重应用性和趣味性，以达到提高学生分析解决实际问题的能力。在教材的编写中，主要概念以定义的形式给出，主要结论以定理的形式给出，帮助学生抓住重点。每章的习题中编排了较多的选择题和填空题，希望学生通过做这些题目，加深对概率论与数理统计的基本概念理解，掌握概率论与数理统计的基本方法和基本技能，提高学生的动手能力。在数理统计部分，扩展内容用“*”标注，可供学生自学或选学，相应的习题也用“*”标注。数理统计部分每章最后配了一个案例分析，以一篇小论文的形式呈现，供学生阅读，借以帮助学生提高分析解决实际问题的能力。

本书第三部分基于 EXCEL 环境安排了常用随机变量分布（包括二项分布、泊松分布、指数分布、正态分布、卡方分布、 t 分布和 F 分布）、参数的区间估计、假设检验、方差分析和回归分析实验。其目的是要求学生能在 EXCEL 环境下，正确选用相应统计工具进行实验，能够正确解读实验结果，提高学生解决实际问题的能力，并为学习应用其他统计软件打下基础。

本书共十五章，其中第一章随机事件及其概率、第二章随机变量及其分布、第三章多维随机变量及其分布、第四章随机变量的数字特征、第五章常

用分布与极限理论、第六章数理统计基础和第三部分基于 EXCEL 的概率统计实验由黄龙生编写，第七章参数估计由黄敏编写，第八章假设检验由宋红凤编写，第九章方差分析由许芳忠编写，第十章回归分析由夏慧珠编写，全书由黄龙生统稿。

本书在编写的过程中参考了大量文献，参考文献中未能尽数列出，在此谨向所有参考文献的编著者表示感谢和敬意。由于编者水平有限，书中不妥之处在所难免，恳请读者批评指正。

编 者
2017 年 10 月

目录

前言

第一部分 概 率 论

第一章 随机事件及其概率	1
第一节 随机事件与概率的基本概念	1
第二节 概率的计算	8
第三节 条件概率	12
第四节 随机事件的独立性	18
习题	23
第二章 随机变量及其分布	27
第一节 随机变量及其分布函数	27
第二节 离散型随机变量及其分布	29
第三节 连续型随机变量及其分布	31
第四节 随机变量函数的分布	35
习题	37
第三章 多维随机变量及其分布	42
第一节 多维随机变量及其联合分布	42
第二节 二维离散型随机变量	44
第三节 二维连续型随机变量	46
第四节 随机变量的独立性	48
第五节 二维随机变量函数的分布	51
第六节 条件分布	54
习题	57
第四章 随机变量的数字特征	65
第一节 随机变量的数学期望	65
第二节 随机变量的方差	70
第三节 协方差与相关系数	72
习题	79

第五章 常用分布与极限理论 83

第一节 常用离散型随机变量的分布 83

第二节 常用连续型随机变量的分布 87

第三节 极限理论 96

习题 101

第二部分 数 理 统 计

第六章 数理统计基础 106

第一节 数理统计的基本概念 107

第二节 数理统计中常用的三大分布 111

第三节 抽样分布 115

* 第四节 数据整理 119

* 案例：全国各地区生产总值的描述性分析 125

习题 130

第七章 参数估计 135

第一节 参数的点估计 135

第二节 点估计效果的评价标准 142

第三节 参数的区间估计 145

* 案例：有重大科学突破时科学家年龄的估计 155

习题 157

第八章 假设检验 162

第一节 假设检验的基本概念 162

第二节 参数的假设检验 167

第三节 假设检验问题的 P -值 177

* 第四节 正态性检验 180

* 第五节 独立性的列联表检验 184

* 第六节 大样本检验 186

* 案例：大西洋两岸过关机场等级测评 187

习题 190

第九章 方差分析 197

第一节 单因素方差分析 197

第二节 双因素方差分析 202

* 案例：全国各地区农村居民恩格尔系数差异分析 212

习题 215

第十章 回归分析 221

第一节 一元线性回归方程 221

第二节 一元回归方程的应用 232

* 案例：用工业品出厂价格指数（PPI）预测居民消费价格指数（CPI）	239
习题	245

第三部分 基于 EXCEL 的概率统计实验

第十一章 常用随机变量的分布实验	250
实验题	261
第十二章 参数的区间估计实验	262
实验题	267
第十三章 参数的假设检验实验	269
实验题	279
第十四章 方差分析实验	281
实验题	285
第十五章 回归分析实验	287
实验题	291
附录	292
附录一 模拟试题	292
附录二 检验表	299
参考文献	309

第一部分 概 率 论

第一章 随机事件及其概率

概率论与数理统计是研究和揭示随机现象统计规律性的一门学科。概率论与数理统计的理论和方法，在工业、农业、军事、天文、医学、金融、保险、试验设计等人类活动的各个领域，产生着越来越重要的作用。在理论联系实际方面，可以说概率论与数理统计是当今世界上发展最为迅速也是最为活跃的数学分支之一。概率论是研究随机现象中数量规律的数学分支，是数理统计的理论基础。



第一节 随机事件与概率的基本概念

一、随机试验与随机事件

(一) 随机现象

在自然界和人类社会活动中，人们所观察到的现象大致可分为必然现象和随机现象两类。

定义 在一定条件下，必然出现的现象，即只有一个结果，因而可以事先准确预知的现象，称为**必然现象或确定性现象** (Certain phenomenon)。

例如：

- ◆ 每天早晨太阳从东方升起。
- ◆ 同性电荷相互排斥，异性电荷相互吸引。
- ◆ 在自然状态下，水从高处流向低处。

定义 在一定条件下，人们不能事先准确预知其结果的现象，即在一定条件下可能出现也可能不出现的现象，称为**随机现象** (Random phenomenon)。

随机现象在日常生活中也是广泛存在的。例如：

◆ 向上抛一枚硬币，落地后可能正面朝上也可能反面朝上，就是说，“正面朝上”这个结果可能出现也可能不出现；



◆ 掷一枚骰子,可能出现 1, 2, 3, 4, 5, 6 点,至于将掷出哪一点,也是不能事先准确预知的;

◆ 在股市交易中,某只股票的价格受到国家金融政策、上市公司业绩、股民的炒作行为及其他国家股市的涨跌等许多不确定因素影响,下一个交易日该股票的股价可能上升也可能下跌,而且这只股票的最高价和最低价也不能事先确定;

◆ 在射击比赛中,运动员用同一支步枪向一个靶子射击,打出的环数可能不同;

◆ 在某一条生产线上,使用相同的工艺生产出来的产品寿命也可能会有较大差异等.

虽然随机现象在相同的条件下可能的结果不止一个,且不能事先准确预知将会出现什么样的结果,但是经过长期的、反复的观察和实验,人们逐渐发现了所谓结果“不能事先准确预知”只是对一次或几次观察或试验而言,在相同条件下进行大量重复观察或试验时,试验的结果就会呈现出某种规律性,这就是所谓的统计规律性.

(二) 随机试验

为了研究随机现象的数量规律,需要对随机现象进行一些重复观察或试验.在这里,试验作为一个含义广泛的术语,它可以是各种各样的科学实验,也可以是对自然现象或社会现象进行的观察.例如:

◆ 在一批笔记本电脑中任意抽取一台,检测它的寿命.

◆ 向上抛一枚硬币三次,观察其落地后出现正面的次数.

◆ 记录某市火车站售票处一天内售出的车票数等都是试验.

定义 具有下述三个特点的试验称为**随机试验** (Random experiment), 简称为试验, 用大写英文字母 E 表示.

(1) 可重复性: 试验可以在相同的条件下重复进行.

(2) 可观察性: 每次试验的可能结果不止一个,但事先可以明确知道试验的所有可能结果.

(3) 不确定性: 进行一次试验之前不能确定哪一个结果会出现.

以后本书中所提到的试验均指随机试验.

(三) 样本空间

由于随机试验具有可观察性,因此,虽然事先不能确定试验将会出现哪一个结果,但试验的所有可能的基本结果所构成的集合却是已知的.

定义 将随机试验 E 的每个可能的基本结果称为一个**样本点** (Sampling point), 全体样本点组成的集合称为 E 的**样本空间** (Sampling space), 记为 $\Omega = \{\omega\}$, 其中 ω 表示试验的样本点.

【例 1-1】 设 E_1 : 向上抛掷一枚硬币,观察其落地后正面朝上还是反面朝上,则 $\Omega_1 = \{\text{正面}, \text{反面}\}$;

E_2 : 将一枚硬币连续向上抛掷两次,依次观察其落地后正面朝上还是反面朝上,则 $\Omega_2 = \{\text{正正}, \text{正反}, \text{反正}, \text{反反}\}$;

E_3 : 将一枚硬币连续向上抛掷两次,观察其反面朝上的次数,则 $\Omega_3 = \{0, 1, 2\}$;

E_4 : 记录某市火车站售票处一天内售出的车票数,则 $\Omega_4 = \{0, 1, 2, \dots\}$;

E_5 : 在某型号电脑中任取一台检测其使用寿命,则 $\Omega_5 = \{t | t \geq 0\}$;



写出试验的样本空间，是描述随机现象的基础。值得注意的是：即使是相同的试验，由于研究目的不同，其样本空间也可能不同，如 Ω_2 和 Ω_3 。也就是说，样本空间的样本点取决于随机试验和它的研究目的。

(四) 随机事件

定义 随机试验 E 的样本空间 Ω 的子集称为 E 的**随机事件** (Random event)，简称事件 (Event)。常用大写英文字母 A 、 B 、 C 等表示事件。

◆ 随机现象中的某些基本结果组成的集合就是随机事件。

◆ 任何一个样本点 ω 构成的单点集 $\{\omega\}$ 也都是随机事件，称为**基本事件** (Basic events)。

◆ 任何事件可看成是由基本事件复合而成。

◆ 样本空间 Ω ，称为**必然事件** (Certain event)。因为 Ω 本身也是 Ω 的一个子集，故也是事件，在每次试验中必然会出现 Ω 中的某一样本点，所以在任何一次试验中 Ω 必然会发生，故称其为必然事件。

◆ 空集 $\emptyset$ ，称为**不可能事件** (Impossible event)。空集 $\emptyset$ 也是 Ω 的子集，故也是事件。因为空集不包含任何样本点，在任何一次试验中 $\emptyset$ 都不可能发生，所以称其为不可能事件。

◆ 在一次随机试验中，事件 A 发生，是指当且仅当 A 所包含的某一样本点出现。

二、随机事件间的关系与运算

因为样本空间 Ω 就是全体样本点所组成的集合，随机事件是 Ω 的子集，所以事件之间的关系和运算也可按集合间的关系和运算来处理。为了简化以后的概率计算，下面的讨论总是假定在同一个样本空间 Ω (即同一个随机现象) 中进行，下面来了解事件之间关系和运算所代表的概率意义。

(一) 包含关系

定义 若事件 A 发生必然导致事件 B 发生，则称事件 B **包含** (Inclusion relation) 事件 A ，或事件 A 包含于事件 B ，记为 $B \supseteq A$ 或 $A \subseteq B$ 。包含关系维恩 (Venn) 图见图 1-1。

◆ $A \subseteq B$ ，也就是事件 A 中的每一个样本点都是事件 B 的样本点。

◆ 对于任意事件 A ，必有： $\emptyset \subseteq A \subseteq \Omega$ 。

(二) 相等关系

定义 若事件 A 发生必然导致事件 B 发生，同时事件 B 发生必然导致事件 A 发生，则称事件 A 与 B **相等** (Equivalent relation)，记为 $A=B$ 。

◆ $A=B$ ，也就是事件 A 中的样本点与事件 B 的样本点完全相同，即 $A \subseteq B$ 和 $B \subseteq A$ 同时成立。

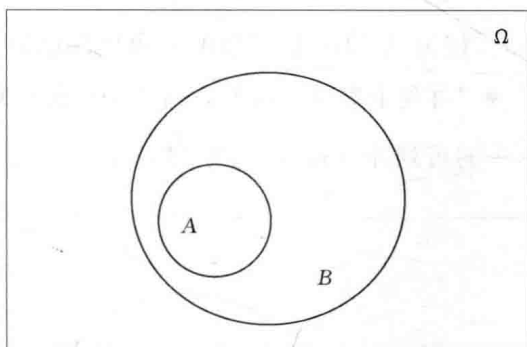


图 1-1 包含关系维恩 (Venn) 图



(三) 互不相容 (互斥) 事件

定义 若事件 A 与 B 不可能同时发生时, 则称事件 A 与事件 B 互不相容 (或互斥) (Incompatible events). 互斥关系维恩图见图 1-2.

- ◆ 事件 A 与事件 B 互斥, 即 $A \cap B = \emptyset$, 事件 A 与 B 没有相同的样本点.
- ◆ 任意两个不同的基本事件是互不相容的.
- ◆ 当 $i \neq j (i, j = 1, 2, \dots)$ 时, $A_i A_j = \emptyset$, 即事件组 $A_1, A_2, \dots, A_n, \dots$ 中任意两个不同事件都互不相容, 则称事件 $A_1, A_2, \dots, A_n, \dots$ 两两互不相容.

(四) 事件的并 (和)

定义 “事件 A 与 B 中至少有一个发生”这一事件称为事件 A 与事件 B 的并 (或和) (Union of events), 记为 $A \cup B$. 并运算维恩 (Venn) 图见图 1-3.

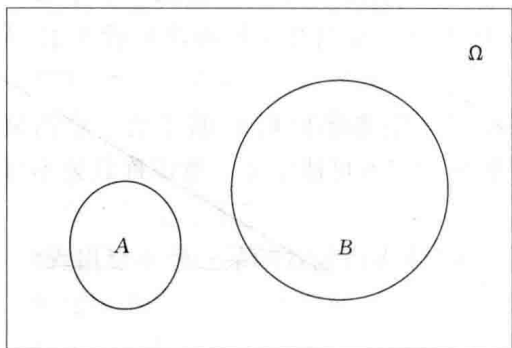


图 1-2 互斥关系维恩 (Venn) 图

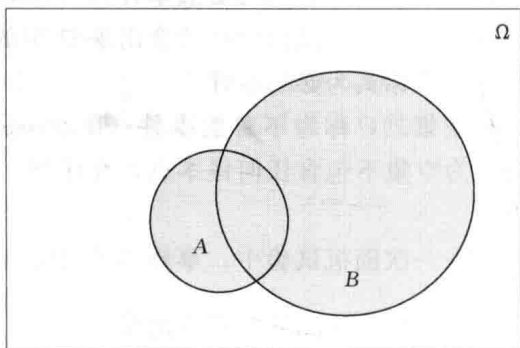


图 1-3 并运算维恩 (Venn) 图

◆ $A \cup B$ 就是由事件 A 和 B 的所有样本点 (相同的只计入一次) 所组成的新事件, 即 $A \cup B = \{\omega | \omega \in A \text{ 或 } \omega \in B\}$.

◆ “事件 $A_1, A_2, \dots, A_n$ 中至少有一个发生”这一事件称为事件 $A_1, A_2, \dots, A_n$ 的并 (和), 记为 $A_1 \cup A_2 \cup \dots \cup A_n$, 也可简记为 $\bigcup_{i=1}^n A_i$.

◆ “可列个事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生”这一事件称为事件 $A_1, A_2, \dots, A_n, \dots$ 的可列并 (和), 记为 $\bigcup_{i=1}^{\infty} A_i$.

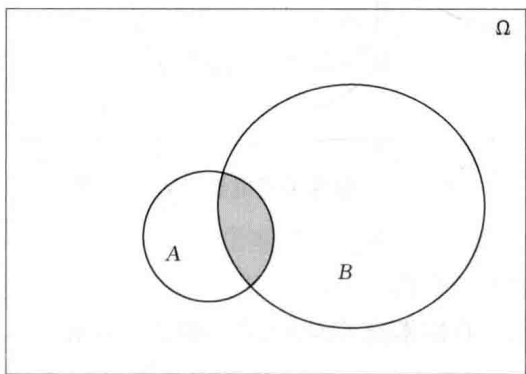


图 1-4 交运算维恩 (Venn) 图

(五) 事件的交 (积)

定义 “事件 A 与 B 同时发生”, 这一事件称为事件 A 与事件 B 的交 (或积) (Product of events), 记为 $A \cap B$, 或简记为 AB , 交运算维恩 (Venn) 图见图 1-4.

◆ $A \cap B$ 就是由事件 A 与 B 中公共的样本点组成的新事件, 这与集合的交集定义完全相同, 即 $A \cap B = \{\omega | \omega \in A \text{ 且 } \omega \in B\}$.

◆ “事件 $A_1, A_2, \dots, A_n$ 同时发生”这一事件称为事件 $A_1, A_2, \dots, A_n$ 的交 (积), 记



为 $A_1 \cap A_2 \cap \cdots \cap A_n$, 或 $A_1 A_2 \cdots A_n$, 也可简记为 $\bigcap_{i=1}^n A_i$.

◆ “可列个事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生” 这一事件称为事件 $A_1, A_2, \cdots, A_n, \cdots$ 的可列交 (积), 记为 $\bigcap_{i=1}^{\infty} A_i$.

(六) 差事件

定义 “事件 A 发生但 B 不发生”, 这一事件称为事件 A 与事件 B 的差事件 (Difference of events), 记为 $A-B$, 差运算维恩 (Venn) 图见图 1-5.

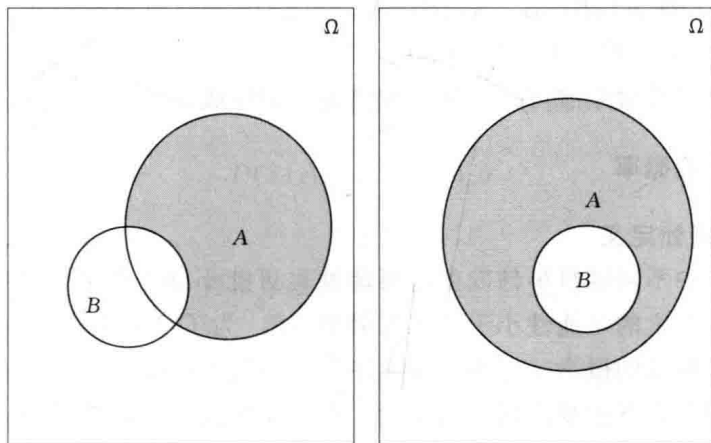


图 1-5 差运算维恩 (Venn) 图

◆ $A-B$ 就是由事件 A 中不属于 B 的样本点组成的新事件, 即

$$A-B = \{\omega \mid \omega \in A \text{ 且 } \omega \notin B\}$$

例如, 掷一枚骰子, $A = \{\text{出现偶数点}\}$, $B = \{\text{出现点数不超过 4}\}$, 则 $A-B = \{6\}$.

(七) 对立事件

定义 “事件 A 不发生” 这一事件称为事件 A 的对立事件 (Opposite events), 记为 $\bar{A}$. 对立事件维恩 (Venn) 图见图 1-6.

◆ $\bar{A}$ 就是由所有 Ω 中不属于事件 A 的样本点组成的新事件.

对立事件也可采用如下定义: 若事件 A 与 B 满足

$$A \cap B = \emptyset, A \cup B = \Omega$$

则称事件 A 与事件 B 互为对立事件, 记为 $\bar{A} = B, \bar{B} = A$.

◆ $\bar{\bar{A}} = A, \bar{\Omega} = \emptyset, \bar{\emptyset} = \Omega$.

◆ $A \cap \bar{A} = \emptyset, A \cup \bar{A} = \Omega, \bar{\bar{A}} = A$.

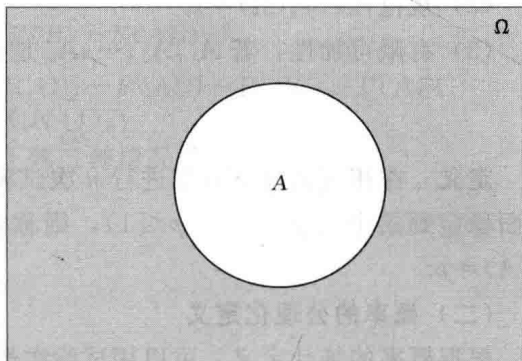


图 1-6 对立事件维恩 (Venn) 图



(八) 事件的运算律

与集合运算一样,事件的运算也满足下列运算规律:

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

(4) 对偶律 (De Morgan 公式):

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

(5) 若 $A \subseteq B$, 则 $A \cup B = B$, $A \cap B = A$.

(6) 事件 A 与 B 的差: $A - B = A \cap \overline{B}$.

上面的运算律对有限个或可列个事件的情况也同样成立.

三、随机事件的概率

(一) 概率的统计定义

在同一个试验中不同随机事件发生的可能性也可能不同. 例如, 在掷骰子的例子中, 显然“出现 6 点”发生的可能性小于“出现偶数点”. 为了度量事件在一次试验中发生的可能性大小, 引入频率的概念, 它描述了事件发生的频繁程度.

定义 在相同条件下重复进行 n 次试验, 在这 n 次试验中, 事件 A 出现的次数 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 (Frequency), 记为 $f_n(A)$, 即

$$f_n(A) = \frac{n_A}{n}$$

容易证明频率具有下述基本性质:

(1) 非负性: $0 \leq f_n(A) \leq 1$.

(2) 规范性: $f_n(\Omega) = 1$.

(3) 有限可加性: 若 $A_1, A_2, \dots, A_n$ 是两两互不相容的事件, 则

$$f_n\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n f_n(A_i)$$

定义 在相同条件下重复进行 n 次试验, 若事件 A 发生的频率随着试验次数 n 的增大而稳定到某个常数 p ($0 \leq p \leq 1$), 则称数值 p 为事件 A 的概率 (Probability), 记作 $P(A) = p$.

(二) 概率的公理化定义

根据概率的统计定义, 可以用试验次数很大时的频率来估计事件的概率. 但是, 在现实生活中某些试验由于成本太高或具有破坏性等原因而不能大量重复进行, 这时不能利用频率来估计概率. 由于概率的统计定义只是一个模糊定义, 不能作为严格的数学定义, 因而存在严重的不足.

历史上还出现过概率的古典定义、概率的几何定义和概率的主观定义, 这些定义只能适应某一类随机现象, 因而不能作为概率的一般定义.

人们通过对概率论中的事件与集合、概率与测度之间联系的研究, 再加上十九世纪末



以来, 数学的各个分支广泛流行着一股公理化潮流, 即把最基本的事实假定为公理, 其他结论均由公理经过演绎导出. 在这种背景下, 1933 年, 苏联数学家柯尔莫哥洛夫 (Kolmogorov) 提出了概率论公理化定义. 这个定义概括了前人所定义的各种概率的共同特征, 又避免了各自的局限性和含糊不清之处, 使概率论成为一门严谨的数学分支, 对概率论的迅速发展起到了积极的作用. 建立在严密的逻辑基础上的概率公理化定义如下.

定义 设 Ω 是随机试验 E 的样本空间, 对于 E 的每一个事件 A , 将其对应于一个实数 $P(A)$, 如果 $P(A)$ 满足下列三个条件, 则称 $P(A)$ 为事件 A 的概率 (Probability).

(1) 非负性: 对任意事件 A , 有 $P(A) \geq 0$.

(2) 规范性: $P(\Omega) = 1$.

(3) 可列可加性: 若 $A_1, A_2, \dots, A_n, \dots$ 是两两互不相容的事件, 则

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

(三) 概率的性质

利用概率定义三条公理, 可以推出概率的另外一些重要性质:

性质 1 $P(\emptyset) = 0$.

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

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$$

性质 3 对任何事件 A , 有 $P(\bar{A}) = 1 - P(A)$.

性质 4 对任意两个事件 A, B 有, $P(A - B) = P(A) - P(AB)$.

◆ 若 $A \supseteq B$, 则 $P(A - B) = P(A) - P(B)$, $P(A) \geq P(B)$.

性质 5 对任意事件 A , 有 $0 \leq P(A) \leq 1$.

性质 6 (加法公式) 对任意事件 A, B, C 有

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

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

【例 1-2】 设 $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{2}$, 就下列三种情况求 $P(B - A)$:

(1) A 与 B 互不相容.

(2) $A \subseteq B$.

(3) $P(AB) = \frac{1}{8}$.

解 (1) 由于 A 与 B 互不相容, 即 $AB = \emptyset$, 所以 $P(B - A) = P(B) - P(AB) = P(B) = \frac{1}{2}$.

(2) $A \subseteq B$, 则有 $P(B - A) = P(B) - P(A) = \frac{1}{4}$.

(3) $P(B - A) = P(B) - P(AB) = \frac{3}{8}$.



第二节 概率的计算

一、古典概率

概率论起源于赌博游戏,因此最先的求概率的问题满足“各个可能结果具有相等或同等可能性”这一假设.例如,在游戏中使用的硬币是均匀的,以保证出现正面和反面的可能性相同;游戏中使用的骰子是均匀的正方体,这样可使得掷出1至6各个点数的可能性相同,从而保证游戏的公平性;一副扑克牌中每一张牌的形状、大小和背面的图案都完全相同,而且在发牌前还要充分地将牌洗匀,使拿到其中每张牌的可能性都相同.

定义 具有以下两条性质的随机试验的概率模型称为古典概型.

- (1) 有限性: 样本空间只含有有限多个样本点;
- (2) 等可能性: 每个基本事件出现的可能性相同.

由于古典概型在产品质量抽样检查等实际问题中有着重要的应用,并且它是最简单的一类随机试验,对它的讨论和研究有助于直观地理解许多概率论中的基本概念,因此在概率论中古典概型占有相当重要的地位.

定理 如果古典概型的样本空间 Ω 包含 n_{Ω} 个样本点,当某个随机事件 A 中所包含的样本点个数为 n_A 时,事件 A 发生的概率就是

$$P(A) = \frac{n_A}{n_{\Omega}} = \frac{A \text{ 包含的样本点数}}{\Omega \text{ 中的样本点总数}}$$

证明 设 $\Omega = \{\omega_i | i=1, 2, \dots, n_{\Omega}\}$, 又记每个基本事件 $A_i = \{\omega_i\}$, $i=1, 2, \dots, n_{\Omega}$, 由古典概型的等可能性易知: $P(A_i) = \frac{1}{n_{\Omega}}$; 又设 $A = \bigcup_{i=1}^{n_A} A_i$, 则

$$P(A) = P\left(\bigcup_{i=1}^{n_A} A_i\right) = \sum_{i=1}^{n_A} P(A_i) = \frac{n_A}{n_{\Omega}}$$

以上确定事件概率的方法称为**古典方法**,这种确定事件概率的方法曾是概率论发展初期的主要方法,故所求的概率又称为**古典概率** (Classical probability).

【例 1-3】 在 1 至 2000 的整数中随机地取一个数,问取到的整数既不能被 6 整除,又不能被 8 整除的概率是多少?

解 设 $A = \{\text{取到的数能被 6 整除}\}$, $B = \{\text{取到的数能被 8 整除}\}$. 由于 $333 < \frac{2000}{6} < 334$, $\frac{2000}{8} = 250$, 故得

$$P(A) = \frac{333}{2000}, \quad P(B) = \frac{250}{2000}$$

又由于一个数同时能被 6 与 8 整除,就相当于能被 24 整除,因此由 $83 < \frac{2000}{24} < 84$ 得

$$P(AB) = \frac{83}{2000}$$



因而所求的概率为

$$\begin{aligned} P(\overline{A\overline{B}}) &= P(\overline{A\cup B}) = 1 - P(A\cup B) = 1 - P(A) - P(B) + P(AB) \\ &= 1 - \frac{333}{2000} - \frac{250}{2000} + \frac{83}{2000} = \frac{3}{4} \end{aligned}$$

【例 1-4】 某机构发售 1 万张即开型福利彩票，其中有 5 张是一等奖，假如你买了 10 张彩票，问你能中一等奖的概率有多大？

解 记 $A = \{\text{能中一等奖}\}$, $A_i = \{\text{第 } i \text{ 张能中一等奖}\}$, 显然, $A = A_1 \cup A_2 \cup \cdots \cup A_{10}$. 直接计算 $P(A)$ 比较麻烦, 但 $\overline{A} = \{\text{没有中一等奖}\}$, $P(\overline{A})$ 的计算则比较简单. 由古典概率计算公式有 $P(\overline{A}) = \frac{C_{9995}^{10}}{C_{10000}^{10}}$, 于是

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{C_{9995}^{10}}{C_{10000}^{10}} \approx 0.00499$$

(一) 计数原理

(1) 加法原理: 若完成某件事有 m 类不同方式, 第一类方式有 n_1 种完成方法, 第二类方式有 n_2 种完成方法, $\cdots$, 第 m 类方式有 n_m 种完成方法, 则完成这件事共有 $n_1 + n_2 + \cdots + n_m$ 种方法.

(2) 乘法原理: 若完成某件事必须经过 m 个不同步骤, 第一个步骤有 n_1 种完成方法, 第二个步骤有 n_2 种完成方法, $\cdots$, 第 m 个步骤有 n_m 种完成方法, 则完成这件事共有 $n_1 \times n_2 \times \cdots \times n_m$ 种方法.

(3) 组合: 从 n 个不同元素中任意取出 r ($1 \leq r \leq n$) 个元素并成一组, 叫做从 n 个不同元素中取 r 个元素的组合. 这时只考虑取出的元素, 不管取出元素的先后次序.

组合数: 从 n 个不同元素中取 r ($1 \leq r \leq n$) 个元素的所有组合的个数, 叫做从 n 个不同元素中取 r 个元素的组合数, 记为 C_n^r .

$$C_n^r = \frac{n!}{r!(n-r)!}$$

(二) 利用计数原理计算古典概率

【例 1-5】 用 0, 1, 2, 3, 4, 5 这六个数字排成三位数, 求:

- (1) 没有相同数字的三位数的概率;
- (2) 没有相同数字的三位偶数的概率.

解 设 $A = \{\text{没有相同数字的三位数}\}$, $B = \{\text{表示没有相同数字的三位偶数}\}$, 则样本点总数 $n_\Omega = 5 \times 6 \times 6 = 180$.

(1) 事件 A 包含的样本点数为 $n_A = 5 \times 5 \times 4$, 所以

$$P(A) = \frac{5 \times 5 \times 4}{5 \times 6 \times 6} = \frac{5}{9}$$

(2) 事件 B 包含的样本点数为 $n_B = 4 \times 4 \times 2 + 5 \times 4 = 52$, 则

$$P(B) = \frac{52}{5 \times 6 \times 6} = \frac{13}{45}$$

【例 1-6】 (抽签的公平性) 口袋中有 a 只黑球, b 只白球. 从袋中不放回地一只一只取球, 直到取完袋中的球为止, 求第 k 次 ($1 \leq k \leq a+b$) 取到黑球的概率.