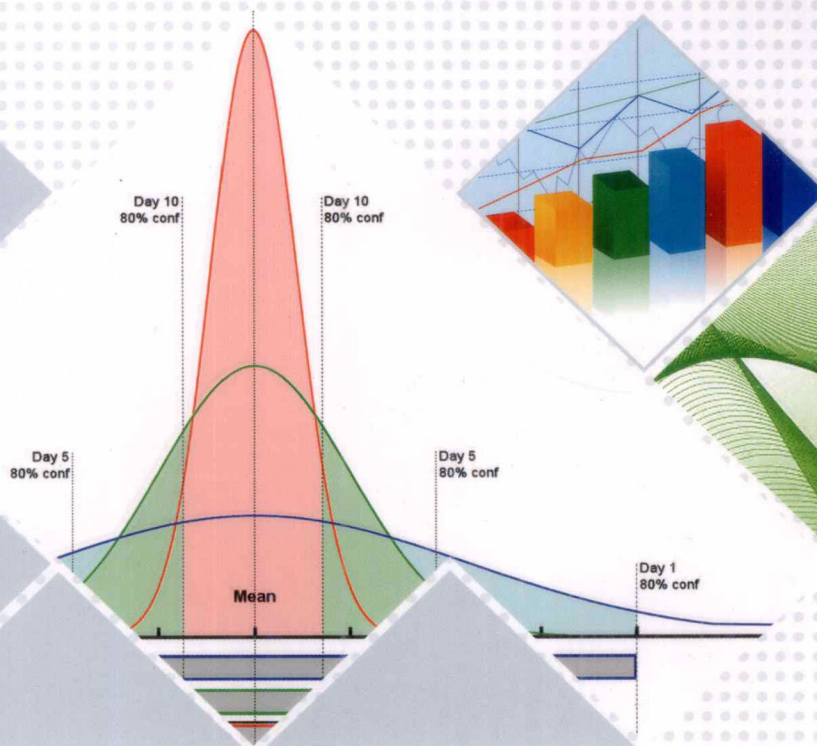




高等教育“十一五”规划教材
公共基础课教材系列

概率论与数理统计

刘焕香 缪春芳 李春华 梁方楚 胡月 等 编著



科学出版社
www.sciencep.com

中国科学院教材建设专家委员会教材建设立项项目

高等教育“十一五”规划教材

公共基础课教材系列

概率论与数理统计

刘焕香 缪春芳 李春华 梁方楚 胡月 等 编著

科学出版社

北 京

内 容 简 介

概率论与数理统计是高等学校经济管理类专业的必修课程,也是学习现代科学技术的重要理论基础。本书以高等学校经济类数学课程的基本要求为依据,在教学实践的基础上编写而成。

全书共分为9章,前4章属于概率论部分的内容,主要介绍概率论的基础知识。第5~8章是数理统计的基本理论和基本统计方法,介绍了参数估计和假设检验,并介绍了方差分析和回归分析。第9章结合现代科学技术的发展趋势,介绍了数理统计实例的计算机实现过程。

本书可以作为高等学校经济管理类(非数学专业)各专业的教材使用,也可作为相关技术人员的参考用书。

图书在版编目(CIP)数据

概率论与数理统计/刘焕香, 缪春芳, 李春华, 梁方楚, 胡月等编著. —北京: 科学出版社, 2009

(中国科学院教材建设专家委员会教材建设立项项目·高等教育“十一五”规划教材·公共基础课教材系列)

ISBN 978-7-03-025428-3

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

中国版本图书馆 CIP 数据核字 (2009) 第 152777 号

责任编辑: 沈力匀 / 责任校对: 赵 燕

责任印制: 吕春珉 / 封面设计: 耕者设计工作室

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>

源海印刷有限责任公司 印刷

科学出版社发行 各地新华书店经销

*

2009年9月第 一 版 开本: 787×1092 1/16

2010年2月第二次印刷 印张: 13 1/4

印数: 4 001—5 500 字数: 314 000

定价: 22.00 元

(如有印装质量问题, 我社负责调换〈环伟〉)

销售部电话: 010-62134988 编辑部电话: 010-62138017 (VP04)

版权所有, 侵权必究

举报电话: 010-64030229; 010-64034315; 13501151303

中国科学院教材建设专家委员会教材建设立项项目

《数学教学改革理论与研究》

专家委员会

主 任 徐宪民

副主任 严从荃 盛宝怀

编 委 (按姓氏笔画排序)

马 万 艾为鸿 兰家诚 郑学良 荆广珠

相丽驰 徐光辉 梅雪锋 韩祥临

前 言

为适应浙江省新课程改革的发展,做好大学数学与高中数学教学内容的衔接,更好地开展概率论与数理统计教学工作,绍兴文理学院等 10 余所院校联合编写了经管类《概率论与数理统计》教材.

全书共分 9 章,第 1 章随机事件及其概率的编写与高中教材衔接,扬长避短、有所侧重;第 2 章强调常用随机变量在社会经济中的应用及其意义;第 3 章中对一些重要概率分析方法和技巧多加归纳和总结,便于掌握;在第 4 章着重分析大数定律和中心极限定理在理论和应用方面的双重价值,体现它们在概率论与数理统计中的桥梁作用;第 5 章强调正态分布抽样定理在统计推断中的重要地位,并贯穿整个数理统计部分;第 6 章参数估计、第 7 章假设检验中始终突出“以样本推断总体”的思想,例题多选取与金融、经济有关的案例,使学生体会数理统计在不同领域中的应用;考虑到经济管理学科的实际应用的需要,在第 8 章对方差分析和回归分析做了简单介绍;第 9 章介绍 Excel 在统计中的简单应用.

本书体现以下特色:

(1) 结合现代计算机技术的使用,在最后一章中介绍如何使用 Excel 软件来实现参数估计、假设检验、方差分析、回归分析等,其操作简单易学、方便实用.

(2) 在习题的编排方面,遵循循序渐进的原则,除第 8、9 章外,每小节后配备练习题,每章编有总习题 A 和 B 两套题目,除了注重基本方法和理论,还列出一些综合性、技巧性强的题目,使学生得到全方位多层次的训练.

本书前 7 章需 45 学时左右,第 8、9 章可根据情况选学. 各章执笔人分别是:第 1 章李春华(浙江万里学院),第 2 章蓝森华(丽水学院),第 3 章王聚丰(宁波理工大学),第 4 章梁方楚(宁波工程技术学院),第 5 章、第 6 章刘焕香(绍兴文理学院),第 7 章胡月(浙江科技学院),第 8 章由绍兴文理学院的缪春芳老师和宁波工程技术学院的梁方楚老师联合编写,第 9 章刘焕香(绍兴文理学院). 本书的统稿和审定工作由绍兴文理学院的刘焕香老师和缪春芳老师共同完成. 本书还得到了绍兴文理学院许多同仁的支持和帮助,在此我们表示衷心的感谢.

由于编者水平有限,书中一定存在不妥之处,诚恳地希望读者批评指正.

目 录

第 1 章 随机事件及其概率	1
§ 1.1 随机事件	1
1.1.1 随机试验与样本空间	1
1.1.2 事件间的关系与运算	2
§ 1.2 随机事件的概率	5
1.2.1 频率	5
1.2.2 概率	6
1.2.3 古典概型	8
§ 1.3 条件概率	11
1.3.1 条件概率	12
1.3.2 乘法公式	14
1.3.3 全概率公式和贝叶斯 (Bayes) 公式	15
§ 1.4 事件的独立性	17
1.4.1 两个事件的独立性	17
1.4.2 多个事件的独立性	19
1.4.3 伯努利 (Bernoulli) 概型	20
总习题 1	22
第 2 章 随机变量及其分布	25
§ 2.1 随机变量	25
§ 2.2 离散型随机变量	25
2.2.1 离散型随机变量及其分布律	25
2.2.2 常见的离散型随机变量	27
§ 2.3 随机变量的分布函数	29
§ 2.4 连续型随机变量	31
2.4.1 连续型随机变量及其概率密度	31
2.4.2 常用的连续型随机变量及其分布	33
§ 2.5 随机变量函数的分布	39
2.5.1 离散型随机变量函数的分布	39
2.5.2 连续型随机变量函数的分布	40
总习题 2	44
第 3 章 多维随机变量及其分布	47
§ 3.1 二维随机变量及其分布	47
3.1.1 二维随机变量及其联合分布函数	47

3.1.2 二维离散型随机变量	48
3.1.3 二维连续型随机变量	50
§ 3.2 边缘分布	53
3.2.1 边缘分布律	53
3.2.2 边缘分布函数	55
3.2.3 边缘概率密度函数	56
§ 3.3 随机变量的独立性	58
3.3.1 二维随机变量的独立性	58
3.3.2 多维随机变量的独立性	61
§ 3.4 多维随机变量函数的分布	62
总习题 3	67
第 4 章 随机变量的数字特征	73
§ 4.1 数学期望	73
4.1.1 离散型随机变量数学期望的定义	73
4.1.2 连续型随机变量数学期望的定义	74
4.1.3 随机变量函数的数学期望	75
4.1.4 数学期望的性质	77
§ 4.2 方差	79
4.2.1 方差的定义	79
4.2.2 方差的性质	81
§ 4.3 协方差与相关系数	83
4.3.1 协方差	83
4.3.2 矩	86
§ 4.4 大数定律和中心极限定理	86
4.4.1 大数定律	86
4.4.2 几个常用的中心极限定理	88
总习题 4	91
第 5 章 数理统计的基本概念	95
§ 5.1 总体与样本	95
5.1.1 总体与样本	95
5.1.2 统计量	96
§ 5.2 抽样分布	98
5.2.1 三大分布	99
5.2.2 抽样分布定理	102
总习题 5	104
第 6 章 参数估计	107
§ 6.1 点估计	107
6.1.1 矩估计法	107

6.1.2 极大似然估计法	109
§ 6.2 估计量的评选标准	113
6.2.1 无偏性	113
6.2.2 有效性	114
6.2.3 一致性	115
§ 6.3 区间估计	116
6.3.1 区间估计的概念	116
6.3.2 单正态总体参数的区间估计	116
6.3.3 双正态总体参数的区间估计	119
6.3.4 单侧置信区间	122
总习题 6	124
第 7 章 假设检验	127
§ 7.1 假设检验概述	127
7.1.1 假设检验问题	127
7.1.2 假设检验的理论依据	128
7.1.3 假设检验的一般步骤	128
7.1.4 检验的两类错误	130
§ 7.2 单正态总体参数的假设检验	130
7.2.1 单正态总体均值的假设检验	130
7.2.2 方差 σ^2 的假设检验—— χ^2 检验法	134
§ 7.3 双正态总体参数的假设检验	136
7.3.1 双正态总体均值差的假设检验	136
7.3.2 双正态总体方差比的假设检验——F 检验法	139
总习题 7	142
第 8 章 方差分析及回归分析	145
§ 8.1 单因素方差分析	145
8.1.1 方差分析模型及其假设条件	145
8.1.2 平方和的分解	147
8.1.3 假设检验问题的拒绝域	148
8.1.4 对未知参数的估计	150
§ 8.2 一元线性回归模型及其参数估计	151
8.2.1 回归模型	151
8.2.2 一元线性回归模型	152
总习题 8	157
第 9 章 Excel 在统计中的应用	159
§ 9.1 σ^2 未知时对总体均值 μ 的区间估计	159
§ 9.2 σ^2 未知时对总体均值 μ 的假设检验	160
§ 9.3 方差 σ_1^2, σ_2^2 未知但相等时对两正态总体均值差 $\mu_1 - \mu_2$ 的检验	162

§ 9.4 双正态总体方差的检验	164
§ 9.5 单因素方差分析	165
§ 9.6 一元线性回归	167
附录	170
附表 1 几种常用的概率分布	170
附表 2 泊松分布函数表	172
附表 3 标准正态分布表	174
附表 4 t 分布表	175
附表 5 χ^2 分布表	177
附表 6 F 分布表	179
习题答案	186
参考文献	201

第 1 章 随机事件及其概率

我们对抛硬币、掷骰子等这一类的游戏很熟悉,这类游戏的结果均以一种偶然的方式出现,我们把这种在个别试验中其结果呈现出不确定性的现象,叫做随机现象.在客观世界中,随机现象是极为普遍的,例如“明天某强队对弱队比赛可能赢球,也可能输球”,“下周的股市可能上涨,也可能下跌”等.概率论与数理统计就是研究和揭示随机现象统计规律性的一门学科,其理论和方法被广泛地应用于自然科学、社会科学、工程技术和经济管理等诸多领域.本章介绍的随机事件及其概率是概率论中最基本、最重要的概念之一.

§ 1.1 随机事件

1.1.1 随机试验与样本空间

自然界和人类社会生活中普遍存在着两类现象:确定性现象和随机现象.

在一定的条件下必然发生的现象称为**确定性现象**.例如,在一个标准大气压下,水加热到 100°C ,必然会沸腾;向上抛一块石头必然下落;同性电荷相斥,异性电荷相吸;初等函数在其定义域内连续等.

我们把事先能够预知所有可能结果,但在每次试验时不能确定哪种结果将要出现的现象称为**随机现象**.例如,抛一枚硬币或者掷一颗骰子;从一批产品中任取一件产品,此产品可能是正品也可能是次品.

在随机现象中,虽然不能事先预言可能出现的具体结果,但是可以认为“所有可能的结果”是已知的.例如,抛一枚硬币的所有结果只有两个:正面和反面;母兔下崽的只数一定是正整数;110 报警台一天内接到的报警次数一定是非负整数;在某一天内某证券的价格充其量可认为是任意实数等.

为了研究随机现象的数量规律,人们需要进行观察或安排试验.例如,为了研究射击中的规律,可以让射手去反复射击;为了检验一枚骰子是否均匀,可以实际反复投掷等.但有很多随机现象是不能重复的,例如,某场足球赛的输赢是不能重复的,某一天的证券交易也是不能重复的.对这类现象,我们只能进行观察.我们把这类观察或试验统称为统计试验.其中在相同条件下可以重复的统计试验又称为随机试验,简称试验,用 E 表示.概率论主要研究能大量重复的随机现象,但也研究不能重复的随机现象.

随机试验的例子:

E_1 : 抛一枚硬币一次,观察正反面出现的情况(分别用 H 和 T 表示正面和反面);

E_2 : 将一枚硬币连续抛三次,观察正反面出现的情况;

E_3 : 将一枚硬币连续抛三次,观察正面出现的次数;

E_4 : 掷一颗骰子一次,观察出现的点数;

E_5 : 某网站在一分钟内受到的点击次数;

E_6 : 在一批某种型号的节能灯泡中任取一只, 测其寿命;

E_7 : 任选一人, 测量其身高和体重.

为了研究随机现象中的规律, 人们通常将随机试验 E 的所有不同的可能结果组成的集合, 称为 E 的样本空间, 记为 Ω .

一般地, 当试验结果仅有有限多个, 或可列无穷多个时, 样本空间 Ω 中的元素, 即 E 的每一个结果, 称为样本点. 样本点一般用 ω 表示. 例如, 一个盒子中有十个相同的球, 其中 6 个是白色的, 另外 4 个是黑色的, 搅匀后从中任意摸取一球, 在观察取球情况的试验中, 令 $\omega_1 = \{\text{取得白球}\}$, $\omega_2 = \{\text{取得黑球}\}$, 则 $\Omega = \{\omega_1, \omega_2\}$.

【例 1.1】 写出上述各随机试验的样本空间.

解:

$$\Omega_1 = \{H, T\};$$

$$\Omega_2 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\};$$

$$\Omega_3 = \{0, 1, 2, 3\};$$

$$\Omega_4 = \{1, 2, 3, 4, 5, 6\};$$

$$\Omega_5 = \{0, 1, 2, 3, \dots\};$$

$$\Omega_6 = \{t: t \geq 0\}, t \text{ 表示灯泡的寿命};$$

$$\Omega_7 = \{(h, w) | h_1 < h < h_2, w_1 < w < w_2\}, h \text{ 和 } w \text{ 分别表示身高和体重}.$$

应该注意的是, 试验 E_2, E_3 的过程都是将一枚硬币连续抛三次, 但由于试验的目的不一样, 样本空间 Ω_2 和 Ω_3 截然不同, 这说明试验的目的决定试验所对应的样本空间.

进行随机试验时, 人们常关心的往往是满足某种条件的样本点所组成的集合. 例如, 若规定节能灯泡的寿命超过 2500 小时为合格品, 则在试验 E_6 中我们关心的是节能灯泡的寿命是否大于 2500 小时, 满足这一条件的样本点组成 Ω_6 的一个子集 $A = \{t: t > 2500\}$. 我们称 A 为试验 E_6 的一个随机事件.

一般地, 称试验 E 的样本空间 Ω 的子集为 E 的随机事件, 简称为事件. 随机事件通常用字母 A, B, C 等表示. 设 A 是一个事件, 当且仅当试验中出现的样本点 $\omega \in A$ 时, 称事件 A 在该次试验中发生. 例如, 在 E_6 中, 若测出节能灯泡的寿命 $t = 2600$ 小时, 则事件 $A = \{t: t > 2500\}$ 即“节能灯泡为合格品”这一事件在该次试验中发生; 同样, 若测出节能灯泡的寿命 $t = 1800$ 小时, 则在该次试验中事件 A 没有发生.

显然, 要判断一个事件是否在一次试验中发生, 只有当该次试验有了结果以后才知道.

由 Ω 中的每一个样本点组成的单点集称为基本事件. 例如, 试验 E_1 有两个基本事件 $\{H\}$ 和 $\{T\}$, 试验 E_3 有四个基本事件 $\{0\}, \{1\}, \{2\}, \{3\}$.

样本空间 Ω 有两个特殊子集, 一个是 Ω 本身, 由于它包含了试验的所有可能的结果, 所以在每一次试验中它总是发生, 称为必然事件. 另一个子集是空集 $\emptyset$, 它不包含任何样本点, 因此在每一次试验中都不发生, 称为不可能事件.

1.1.2 事件间的关系与运算

事件是一个集合, 因而事件间的关系与运算可以按照集合之间的关系与运算来处理.

要注意的是,我们只是借助集合的符号,需要清楚它们在概率论中的意义.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,给出它们在概率论中的含义.

设随机事件 E 的样本空间为 Ω , $A, B, A_k (k=1, 2, \dots)$ 是 Ω 的子集,即随机事件.

(1)若事件 A 发生必然导致事件 B 发生,则称事件 A 包含于事件 B ,或称事件 A 是 B 的子事件,记作 $A \subset B$.

若 $A \subset B$ 且 $B \subset A$,即 $A=B$,称事件 A 与事件 B 相等.

(2)事件“事件 A 与事件 B 至少有一个发生”,称为事件 A 与事件 B 的和事件,记作 $A \cup B$.

$$A \cup B = \{\omega; \omega \in A \text{ 或 } \omega \in B\}$$

例如,在试验 E_2 中,若事件 A 表示“三次都出现正面”,即 $A=\{HHH\}$,事件 B 表示“三次都出现反面”,即 $B=\{TTT\}$,那么事件 $A \cup B$ 表示“三次出现同一面”,即 $A \cup B = \{HHH, TTT\}$.

类似地,称 $\bigcup_{k=1}^n A_k$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的和事件;称 $\bigcup_{k=1}^{\infty} A_k$ 为可列个事件 $A_1, A_2, \dots, A_n, \dots$ 的和事件.

(3)事件“事件 A 与事件 B 同时发生”,称为事件 A 与事件 B 的积事件,记作 $A \cap B$,可简记为 AB .

$$AB = \{\omega; \omega \in A \text{ 且 } \omega \in B\}$$

例如,某输油管长 100km,事件 A 表示“前 50km 油管正常工作”,事件 B 表示“后 50km 油管正常工作”,那么 $A \cap B$ 表示“整个输油管正常工作”.

类似地,称 $\bigcap_{k=1}^n A_k$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的积事件;称 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件 $A_1, A_2, \dots, A_n, \dots$ 的积事件.

(4)事件“事件 A 发生而事件 B 不发生”,称为事件 A 与事件 B 的差事件,记作 $A-B$.

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

例如,在试验 E_2 中,若事件 A 表示“至少两次出现正面”,事件 B 表示“三次出现同一面”,此时

$$A = \{HHH, HHT, HTH, THH\}$$

$$B = \{HHH, TTT\}$$

则

$$A-B = \{HHT, HTH, THH\}$$

(5)若事件 A 与事件 B 不能同时发生,则称事件 A 与事件 B 是互斥事件或互不相容事件,记作 $AB=\emptyset$.

例如,基本事件是两两互不相容的.

(6)若 $A \cup B = \Omega$ 且 $AB = \emptyset$,则称事件 A 与事件 B 是对立事件或互逆事件,事件 A 的对立事件记作 $\bar{A}$. 其含义是:对每次试验而言,事件 $A, \bar{A}$ 中必有一个且仅有一个发生.

易见,事件的运算满足如下基本关系:

① 若 $A \subset B$,则 $A \cup B = B, AB = A$;

$$\textcircled{2} A-B=A\bar{B}=A-AB, A\cup B=A\cup(B-A);$$

$$\textcircled{3} A\bar{A}=\emptyset, A\cup\bar{A}=\Omega, \bar{A}=\Omega-A.$$

事件的关系与运算可用图 1.1 中的图形更加清楚地表示出来.

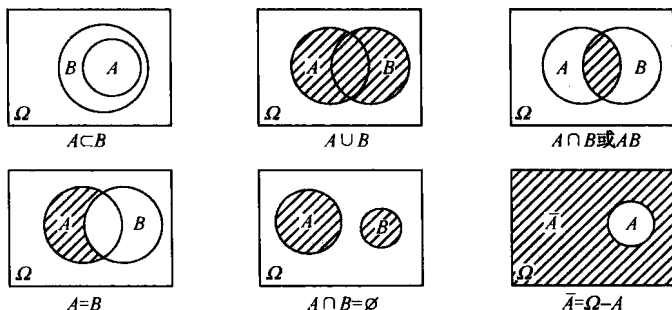


图 1.1

与集合的运算一样,事件间的运算满足下述运算律.

$$(1) \text{ 交换律: } A \cup B = B \cup A; A \cap B = B \cap A.$$

$$(2) \text{ 结合律: } (A \cup B) \cup C = A \cup (B \cup C);$$

$$(A \cap B) \cap C = A \cap (B \cap C).$$

$$(3) \text{ 分配律: } A \cup (B \cap C) = (A \cup B) \cap (A \cup C);$$

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

$$(4) \text{ 对偶律: } \overline{A \cup B} = \bar{A} \cap \bar{B}; \overline{A \cap B} = \bar{A} \cup \bar{B}.$$

这些运算规律可以推广到有限个或可列个事件的情形.

注意以下符号在概率论中的意义.

$\overline{A \cup B}$: 事件 A 与事件 B 一个都不发生;

$\bar{A} \cap \bar{B}$: 事件 A 不发生且事件 B 也不发生;

$\overline{A \cap B}$: 事件 A 与事件 B 不同时发生;

$\bar{A} \cup \bar{B}$: 事件 A 与事件 B 至少有一个不发生.

可见 $\bar{A} \cap \bar{B}$ 与 $\overline{A \cap B}$ 不是同一个事件.

【例 1.2】 某人连续购买体育彩票,事件 A, B, C 分别表示其第一、二、三次所买的彩票中奖,则事件:

(1) “第三次未中奖”可以表示成 $\bar{C}$;

(2) “第三次才中奖”可以表示成 $\bar{A}\bar{B}C$;

(3) “恰有一次中奖”可以表示成 $A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C$;

(4) “至少有一次中奖”可以表示成 $A \cup B \cup C$;

(5) “不止一次中奖”可以表示成

$$AB \cup AC \cup BC = A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C \cup ABC$$

(6) “至多中奖二次”可以表示成 $\overline{ABC}$.

注 通过上面的例子可以看出,用其他事件的运算来表示一个事件,方法往往不唯一,读者应学会用不同方式表达同一事件,特别在解决具体问题时,要根据需要选择一种

恰当表示方法.

习题 1.1

- 写出下列随机试验的样本空间.
 - 抛三枚硬币,观察正反面出现的情况;
 - 连续抛一枚硬币,直到出现正面为止,记录试验结果;
 - 掷两枚骰子,观察出现的点数;
 - 某城市一天内的用电量.
- 请指出以下事件 A 与 B 之间的关系.
 - 检查两件产品,记事件 A = “至少有一件不合格品”, B = “两次检查结果不同”.
 - 设 T 表示轴承寿命,记事件 $A = \{T < 5000h\}$, $B = \{T > 8000h\}$.
- 袋中有三个球编号为 1、2、3,从中任意摸出一球,观察其号码,记事件 A = “球的号码小于 3”, B = “球的号码为奇数”, C = “球的号码为 3”. 试问:
 - 试验的样本空间是什么?
 - A 与 B , A 与 C , B 与 C 是否互不相容?
 - A, B, C 对立事件是什么?
 - A 与 B 的和事件,积事件,差事件各是什么?
- 某工人加工了三个零件,设 A_i 表示事件“加工的第 i 个零件是合格品”, $i = 1, 2, 3$, 试用 A_1, A_2, A_3 表示下面的事件.
 - 只有第一个零件是合格品;
 - 只有一个零件是合格品;
 - 至少有一个零件是合格品;
 - 最多有一个零件是合格品.

§ 1.2 随机事件的概率

除必然事件和不可能事件外,任一事件在一次试验中可能发生,也可能不发生. 我们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,商业保险机构为获得较大利润,就必须研究个别意外事件发生的可能性大小,从而计算保险费和赔偿金的多少. 为此,先引入频率,它描述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.

1.2.1 频率

定义 1.1 事件 A 在 n 次重复试验中出现 n_A 次,则比值 $\frac{n_A}{n}$ 称为事件 A 在 n 次重复试验中出现的频率,记作 $f_n(A)$,即

$$f_n(A) = \frac{n_A}{n} \quad (1.1)$$

由频率的定义易见其有下列基本性质.

(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: 若事件 $A_1, A_2, \dots, A_k$ 两两互不相容, 即 $A_i A_j = \emptyset, i, j = 1, 2, \dots, k$ 且 $i \neq j$, 则

$$f_n\left(\bigcup_{i=1}^k A_i\right) = \sum_{i=1}^k f_n(A_i) \quad (1.2)$$

事实上, 频率还有稳定性, 即在大量重复试验中任一事件的频率都呈现出明显的稳定性. 也就是说, 当试验的次数 n 逐渐增大时, 随机事件 A 的频率 $f_n(A)$ 常在某个确定的常数附近摆动, 而逐渐稳定于这个常数. 这个常数是客观存在的.

比如抛硬币试验中, “正面朝上”这一随机事件 A 的频率 $f_n(A)$ 稳定在数字 0.5 的附近. 历史上不少著名学者做过抛掷硬币试验, 得到的数据如表 1.1 所示.

表 1.1

试验者	n	n_A	$f_n(A)$
德·摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
K. 皮尔逊	12000	6019	0.5016
K. 皮尔逊	24000	12012	0.5005
罗曼诺夫斯基	80640	39699	0.4923

关于频率的稳定性, 生活中也有许多有趣的例子. 例如, 衣服和鞋子总在同样部位以相似的方式破损, 下雨时地面各处总是差不多同时淋湿等.

通常将频率的稳定值称为事件 A 的概率, 记作 $P(A)$. 这样引入的定义称为概率的统计定义. 由此定义可知, 在抛硬币的试验中, “正面朝上”这一随机事件 A 发生的概率是 0.5, 这与人们的直观判断是一致的.

但是, 实际生活中有些试验不可重复进行, 无法计算事件发生的频率, 即使对可重复进行的试验, 也不可能对每一个事件做大量的试验, 然后求出事件的频率, 用以表征事件发生的可能性的. 因此, 需要引出一个能够揭示概率本质属性的定义.

1.2.2 概率

任何一个数学概念都是对现实世界的抽象, 这种抽象使得其具有广泛的适应性, 并成为进一步数学推理的基础. 1933 年苏联数学家柯尔莫哥洛夫综合已有的大量成果, 提出了概率的公理化结构, 第一次将概率论建立在严密的逻辑基础上.

下面给出概率的公理化定义.

定义 1.2 设 E 是随机试验, Ω 是它的样本空间, 对于 E 的每一个事件 A 赋予一实数 $P(A)$, 若集合函数 $P(\cdot)$ 满足条件:

(1) 非负性: $P(A) \geq 0$;

(2) 规范性: $P(\Omega) = 1$;

(3) 可列可加性: 若事件 $A_1, A_2, \dots$ 两两互不相容, 即 $A_i A_j = \emptyset, i, j = 1, 2, \dots$, 且 $i \neq j$, 则有

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots \quad (1.3)$$

则称 $P(A)$ 为事件 A 的概率.

由概率的定义, 可以推出概率的一些重要性质.

性质 1 $P(\emptyset) = 0$;

性质 2 (有限可加性) 若事件 $A_1, A_2, \dots, A_n$ 两两互不相容, 即 $A_i A_j = \emptyset, i, j = 1, 2, \dots, n$, 且 $i \neq j$, 则有

$$P(A_1 \cup A_2 \cup \dots \cup A_{n-1} \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n);$$

性质 3 对于事件 A, B , 若 $A \subset B$, 则有

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

$$P(A) \leq P(B)$$

一般地, 对于任意两个事件 A, B , 有 $P(B - A) = P(B) - P(AB)$;

上式称为概率的减法公式.

性质 4 对于任一事件 $A, 0 \leq P(A) \leq 1$;

性质 5 (逆事件的概率) 对于任一事件 A , 有 $P(\bar{A}) = 1 - P(A)$;

性质 6 (加法公式) 对于任意两事件 A, B , 有

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

性质 1~4 的证明留给读者, 这里仅给出性质 5、6 的证明.

证明(性质 5) 因为 $A \cup \bar{A} = \Omega$, 且 $A\bar{A} = \emptyset$, 由性质 2 可得

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$$

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

证明(性质 6) 因为 $A \cup B = A \cup (B - AB)$, 且 $A(B - AB) = \emptyset, AB \subset B$, 故由性质 2 和性质 3 得

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

性质 6 可以推广到多个事件的情况. 如 A, B, C 是任意三个事件, 则有

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

更一般地, 对于任意 n 个事件 $A_1, A_2, \dots, A_n$, 可以用归纳法证明得到

$$\begin{aligned} P\left(\bigcup_{k=1}^n A_k\right) &= \sum_{k=1}^n P(A_k) - \sum_{1 \leq i < j \leq n} P(A_i A_j) \\ &\quad + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n) \end{aligned} \quad (1.4)$$

【例 1.3】 设事件 A, B 互不相容, 已知 $P(A) = p, P(B) = q$, 求 $P(A \cup B), P(\bar{A}B), P(\bar{A} \cup B), P(\bar{A}\bar{B})$.

解 由于 $AB = \emptyset, \bar{A}B = B - A = B - AB$,

所以

$$P(A \cup B) = P(A) + P(B) = p + q$$

$$P(\bar{A}B) = P(B) - P(AB) = P(B) = q$$

$$P(\bar{A} \cup B) = P(\bar{A}) + P(B) - P(\bar{A}B) = 1 - p$$

由对偶律得

$$P(\bar{A}\bar{B}) = 1 - P(A \cup B) = 1 - p - q$$

【例 1.4】 某市有 A, B, C 三种报纸, 订每种报纸的人数占全体市民人数的 30%, 其中有 10% 的人同时订 A, B 两种报纸, 没有人同时订 A, C 或 B, C 两种报纸. 求从该市任选一人, 他至少订有一种报纸的概率.

解 设 A 表示事件“从该市任选一人, 他订 A 报”, B 表示事件“从该市任选一人, 他订 B 报”, C 表示事件“从该市任选一人, 他订 C 报”, 则

$$P(A) = P(B) = P(C) = 0.30, P(AB) = 0.10, P(AC) = P(BC) = 0$$

由于 $ABC \subset BC$, 而

$$0 \leq P(ABC) \leq P(BC) = 0$$

所以

$$P(ABC) = 0$$

由加法公式, “从该市任选一人, 他至少订有一种报纸”的概率为

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) \\ &= 0.80 \end{aligned}$$

1.2.3 古典概型

我们称具有下列两个特征的随机试验模型为**古典概型**.

- (1) 试验的样本空间是有限集, 即 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$;
- (2) 试验中每个基本事件发生的可能性都相同, 即

$$P(\omega_1) = P(\omega_2) = \dots = P(\omega_n)$$

古典概型又称为**等可能概型**, 是概率论发展初期的主要研究对象, 它比较简单且直观, 又能概括许多实际问题, 有着广泛的应用.

定义 1.3 设古典概型试验 E 的样本空间 Ω 有 n 个样本点, 若事件 A 包含其中的 k 个样本点, 则事件 A 的概率定义为

$$P(A) = \frac{k}{n} = \frac{\text{事件 } A \text{ 包含的基本事件数}}{\Omega \text{ 中的基本事件数}} \quad (1.5)$$

由古典概型的两个特征不难看出上述定义的合理性.

【例 1.5】 将一枚硬币抛掷 3 次, A_1 表示事件“恰有一次出现正面”, A_2 表示事件“至少有一次出现正面”求 $P(A_1), P(A_2)$.

解 $\Omega = \{HHH, THH, HTH, HHT, HTT, THT, TTH, TTT\}$

$$A_1 = \{HTT, THT, TTH\}$$

$$A_2 = \{HTT, THT, TTH, HHT, HTH, THH, HHH\}$$

所以

$$P(A_1) = \frac{3}{8}, P(A_2) = \frac{7}{8}$$

注 本题中, 若考虑出现正面的次数, 则样本空间为 $\Omega = \{0, 1, 2, 3\}$, 各基本事件发生