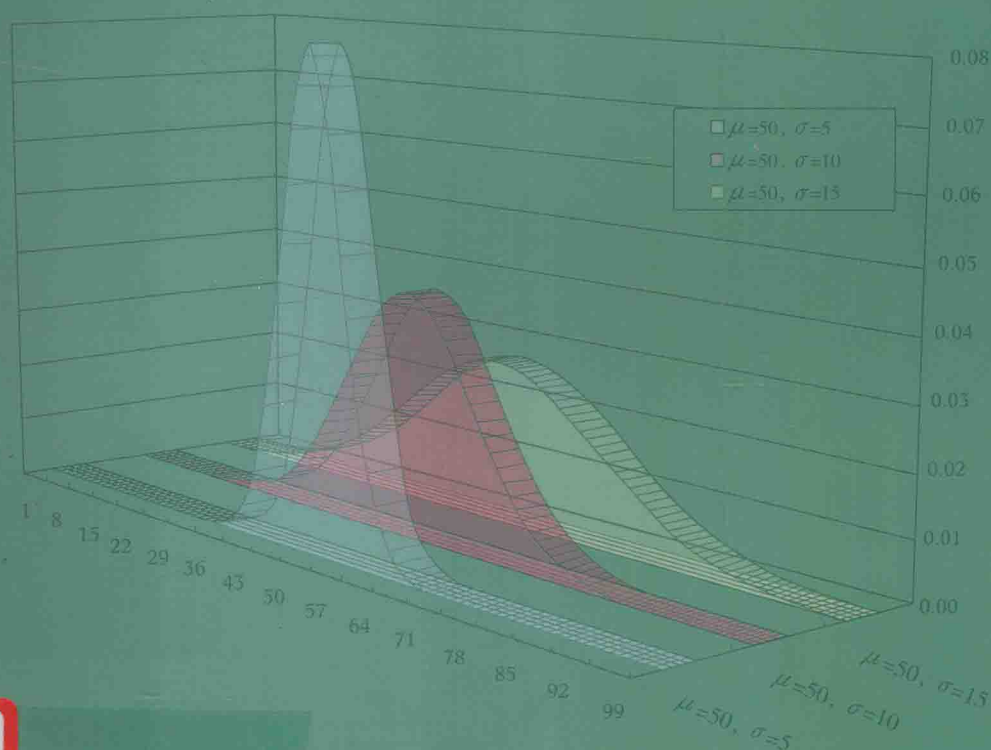


经济应用数学基础（三）

概率论与数理统计

姚孟臣 / 编著

第二版

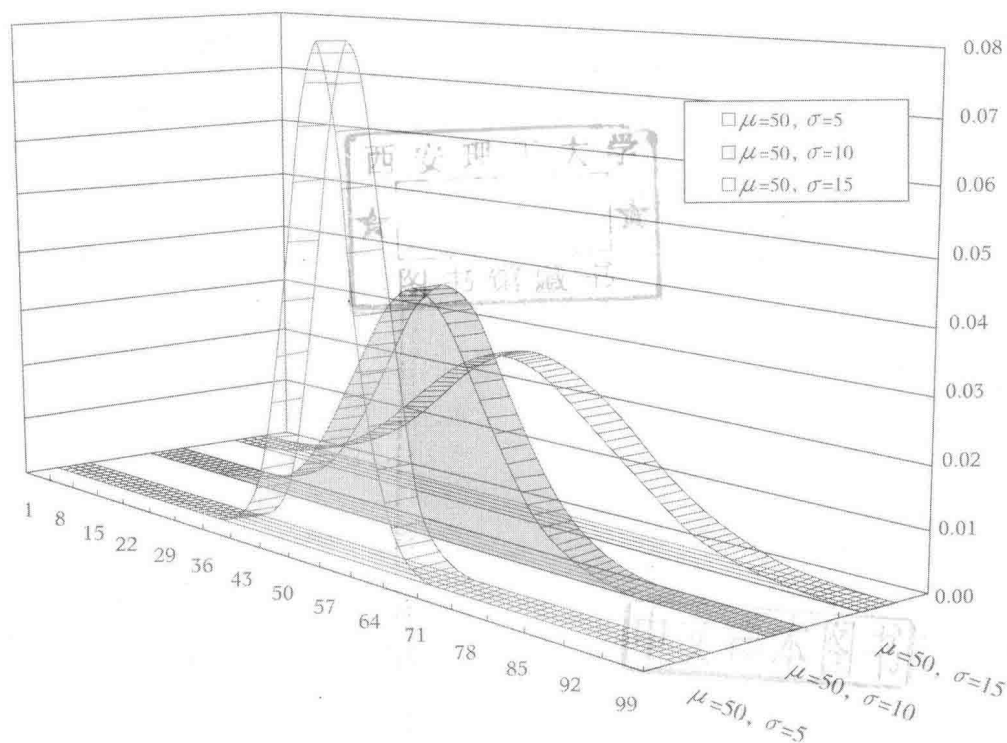


021
218
2016
经济应用数学基础（三）

概率论与数理统计

姚孟臣 / 编著

第二版



中国人民大学出版社

· 北京 ·

图书在版编目(CIP)数据

概率论与数理统计/姚孟臣编著. —2 版. —北京: 中国人民大学出版社, 2016. 6
(经济应用数学基础)
ISBN 978-7-300-22934-8

I. ①概… II. ①姚… 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 (2016) 第 113889 号

经济应用数学基础 (三)
概率论与数理统计 (第二版)
姚孟臣 编著
Gailulun yu Shuli Tongji

出版发行	中国人民大学出版社		
社 址	北京中关村大街 31 号	邮政编码	100080
电 话	010-62511242 (总编室)		010-62511770 (质管部)
	010-82501766 (邮购部)		010-62514148 (门市部)
	010-62515195 (发行公司)		010-62515275 (盗版举报)
网 址	http://www.crup.com.cn		
	http://www.ttrnet.com (人大教研网)		
经 销	新华书店		
印 刷	北京东方圣雅印刷有限公司	版 次	2010 年 6 月第 1 版
规 格	185 mm×260 mm 16 开本		2016 年 6 月第 2 版
印 张	13.5 插页 1	印 次	2016 年 6 月第 1 次印刷
字 数	298 000	定 价	28.00 元

版权所有 侵权必究

印装差错 负责调换



再版前言

《概率论与数理统计》作为经管类公共基础数学教材,它涵盖了经济管理专业有关教学大纲的全部内容 with 基本要求,但是由于各个学校不同专业方向的学生对数学基础知识的掌握存在一定的差异,我们建议,使用时有些内容如方差分析、回归分析等可视教学需要与学时安排略去不讲;有些较难的习题,如(B)中的某些题目则不要求学生掌握.

在本书编写和这次修订过程中,我们参考了有关教材和著作,并且从中摘取了一些例题和习题,书中没有一一注明,再次一并向有关作者致谢!

由于编者水平有限,书中难免有不妥之处,恳请作者批评指正.

编者

2016年4月



目 录

第 1 章 随机事件及其概率	1
§ 1.1 随机事件	1
§ 1.2 概率	6
§ 1.3 条件概率与全概公式	15
§ 1.4 事件的独立性与伯努利概型	20
习题一	25
第 2 章 随机变量及其分布	29
§ 2.1 随机变量与分布函数	29
§ 2.2 离散型随机变量及其分布	32
§ 2.3 连续型随机变量及其分布	38
§ 2.4 二维随机变量	46
§ 2.5 随机变量函数的分布	57
习题二	64
第 3 章 随机变量的数字特征	69
§ 3.1 数学期望	69
§ 3.2 方差	75
§ 3.3 几种常见分布的数学期望与方差	79
§ 3.4 随机变量矩、协方差与相关系数	81
习题三	86
第 4 章 大数定律与中心极限定理	89
§ 4.1 切比雪夫不等式	89
§ 4.2 大数定律	91
§ 4.3 中心极限定理	93
习题四	96

第5章 抽样分布	99
§ 5.1 总体与样本	99
§ 5.2 样本函数与样本分布函数	100
§ 5.3 抽样分布	104
习题五	110
第6章 参数估计	112
§ 6.1 点估计	112
§ 6.2 估计量的评价标准	117
§ 6.3 区间估计	121
§ 6.4 正态总体均值与方差的区间估计	122
§ 6.5 非正态总体参数的区间估计	129
习题六	131
第7章 假设检验	133
§ 7.1 假设检验的基本概念	133
§ 7.2 单个正态总体参数的假设检验	136
§ 7.3 两个正态总体参数的假设检验	143
§ 7.4 非正态总体参数的假设检验	146
§ 7.5 总体分布的假设检验	150
习题七	154
第8章 方差分析	156
§ 8.1 问题的提出	156
§ 8.2 单因素试验方差分析	157
§ 8.3 单因素方差分析举例	163
习题八	165
第9章 回归分析	167
§ 9.1 问题的提出	167
§ 9.2 一元正态线性回归	168
§ 9.3 一元非线性回归简介	175
§ 9.4 多元线性回归	177
§ 9.5 多元回归应用举例	184
习题九	191
附录 常用分布表	193
附表1 泊松分布表	193
附表2 标准正态分布表	195
附表3 χ^2 分布表	196

附表 4 t 分布表	197
附表 5 F 分布表	198
附表 6 检验相关系数的临界值表	202
习题参考答案	203



第 1 章 随机事件及其概率

概率论是从数量上研究随机现象统计规律性的一门数学学科,其应用非常广泛,是科技、管理、经济等领域的工作者必备的数学工具. 本章将向大家介绍概率论中的几个基本概念,随机事件的基本关系与基本运算,以及概率的性质及其计算方法.

§ 1.1 随机事件

(一) 随机现象

在生产实践、科学实验和日常生活中,人们观察到的现象大体可归结为两种类型. 一类是**确定性现象**,即在一定条件下,必然会发生某一种结果或必然不发生某一种结果的现象,例如,在一个标准大气压下,纯净的水加热到 100°C 时必然会沸腾;从 10 件产品(其中 2 件是次品,8 件是正品)中,任意地抽取 3 件进行检验,这三件产品决不会全是次品;向上抛掷一枚硬币必然下落等都是确定性现象. 这类现象的一个共同点是:事先可以断定其结果.

另一类是**随机现象**,即在一定条件下,具有多种可能发生的结果的现象. 例如,从 10 件产品(其中 2 件是次品,8 件是正品)中,任取 1 件,则该产品可能是正品,也可能是次品;向上抛掷一枚硬币,落下以后可能是正面朝上,也可能是反面朝上;新出生的婴儿可能是男性,也可能是女性. 这类现象的一个共同点是:事先不能预言多种可能结果中究竟出现哪一种. 随机现象是偶然性与必然性的辩证统一,其偶然性表现在每一次试验前,都不能准确预言发生哪种结果;其必然性表现在相同条件下多次重复某一个试验时,其各种结果会表现出一定的量的规律性,我们称之为**随机现象的统计规律性**.

概率论就是一门研究随机现象统计规律性的数学分支,它从表面上看起来错综复杂的偶然现象中揭示出潜在的必然性.

(二) 随机试验

为研究随机现象的统计规律性而进行的各种科学实验或对事物某种特征进行的观测

都称为**随机试验**,简称**试验**,用字母 E 表示.例如:

E_1 : 抛一枚均匀硬币,观察它自由下落后正、反面出现的情况.

E_2 : 在相同条件下,连续不断地向同一个目标射击,直到击中为止,记录射击次数.

E_3 : 在一批同型号的灯泡中,任意抽取一只,测试它的使用寿命.

试验 E_1 只有两种可能结果: 出现正面或出现反面,但是抛之前不知道究竟出现哪一面. 对于试验 E_2 , 射击次数可以为 $1, 2, \dots$, 因此试验的所有可能结果是全体正整数, 在击中目标前, 究竟需要射击多少次不能事先确定. 对于试验 E_3 , 灯泡的寿命(以小时计)是一个非负的实数, 在测试结束前不能确定它的寿命有多长. 这些试验都具有下列三个特性:

(1) 在相同的条件下, 试验可以或原则上可以重复进行, 即重复性;

(2) 每次试验的结果具有多种可能性, 但是在试验之前可以明确一切可能出现的基本结果, 即明确性;

(3) 每次试验之前不能准确地预言这次试验会出现哪个结果, 但可以确定每次试验总会出现这些可能结果中的某一个, 即随机性.

(三) 随机事件

随机试验的每一种可能的结果称为**随机事件**, 简称**事件**, 用大写拉丁字母 $A, B, C, \dots$ 表示, 必要时可加上下标. 例如 A = “正面向上”, B = “抽到合格品”, C = “灯泡的使用寿命低于 1000 小时”等都是随机事件. 在一定的研究范围内, 不能再分解的最简单的随机事件称为**基本事件**, 用 ω 表示.

例 1 设试验 E 为掷一颗骰子, 观察其出现的点数. 在这个试验中, 记事件 ω_n = “出现点数 n ”, $n = 1, 2, 3, 4, 5, 6$. 显然, $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$ 都是基本事件. 除此之外, 若记 A = “出现奇数点”, B = “出现被 3 整除的点”, 则 A, B 也都是随机事件, 其中事件 A 是由 $\omega_1, \omega_3, \omega_5$ 这三个基本事件组成的, 事件 B 是由 ω_3, ω_6 这两个基本事件组成的.

由两个或两个以上基本事件组成的事件称为**复合事件**. 例 1 中, 事件 A 和 B 都是复合事件.

在一定条件下, 必然发生的事件称为**必然事件**, 用字母 U 或符号 Ω 表示. 例如, “掷一颗骰子, 出现点数大于零”, “抛一枚硬币, 落下后, 正面向上或反面向上至少有一个发生”, 都是必然事件. 在一定条件下, 必然不发生的事件称为**不可能事件**, 用字母 V 或符号 $\emptyset$ 表示. 例如, “掷一颗骰子, 出现点数大于 7”, “抛一枚硬币, 落下后, 正面向上和反面向上同时发生”, 都是不可能事件.

需要指出的是: 必然事件与不可能事件是每次试验之前都可以准确预知的, 因此它们不是随机事件, 但是为了讨论问题方便, 把它们都看成是特殊的随机事件, 即作为随机现象的两个极端情况.

运用点集的概念研究事件, 将使问题变得更加直观且容易理解.

将随机试验 E 的全部基本事件构成的集合称为**基本事件空间**或**样本空间**, 记为 Ω , Ω 中的元素就是 E 的基本事件, 也称为**样本点**; 由一些基本事件构成的复合事件用由这些基本事件对应的样本点所构成的集合表示, 它是样本空间的一个子集. 这样, 样本空间

作为一个事件是必然事件. 一次试验中某随机事件发生当且仅当该集合所包含的某一个样本点在试验中出现. 不可能事件在每次试验中都不会发生, 因此, 空集 $\emptyset$ 表示不可能事件.

例如, 在例1的试验 E 中, 样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$, 基本事件 $A_i = \{i\}$, $i=1, 2, \dots, 6$. $A = \{1, 3, 5\}$ 是样本点的集合, 事件 A 发生, 即1, 3, 5这三个样本点中至少有一个在试验中发生.

(四) 事件的关系和运算

任何一个随机试验中总有许多随机事件, 其中有些比较简单, 有些则相当复杂. 为了从较简单事件发生的规律中寻求较复杂事件发生的规律, 需要研究同一试验的各种事件之间的关系和运算.

1. 事件的包含关系

若事件 A 发生必然导致事件 B 发生, 即 A 中的每一个样本点都包含在 B 中, 则称事件 B 包含事件 A , 或称事件 A 含于事件 B , 记作 $B \supset A$ 或 $A \subset B$. 对任意事件 A , 有 $\emptyset \subset A \subset \Omega$.

我们用维恩(Venn)图对这种关系给出直观的说明. 图1—1中的长方形表示样本空间 Ω , 长方形内的每一点表示样本点, 圆 A 和圆 B 分别表示事件 A 和事件 B . 如图1—1所示, 圆 A 在圆 B 的里面表示事件 B 包含事件 A .

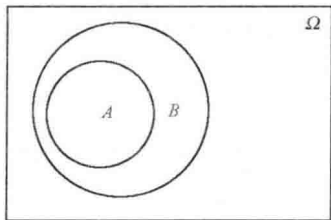


图 1—1

在例1中, 设 $A = \{\omega_2\}$, $B = \{\text{出现偶数点}\}$, 则 $B \supset A$.

2. 事件的相等关系

若事件 A 包含事件 B , 且事件 B 包含事件 A , 则称事件 A 与事件 B 相等, 即 A 与 B 所含的样本点完全相同. 记作 $A = B$.

3. 事件的并(和)

设 A, B 为两个事件. 我们把至少属于 A 或 B 中一个的所有样本点构成的集合称为事件 A 与 B 的并或和, 记为 $A \cup B$ 或 $A + B$. 这就是说, 事件 $A \cup B$ 表示在一次试验中, 事件 A 与 B 至少有一个发生. 图1—2中的阴影部分表示 $A \cup B$.

事件和的概念可以推广到有限个或可列无穷多个事件.

事件 $A_1, A_2, \dots, A_n$ 中至少有一个发生的事件 A 称为这 n 个事件 $A_i (i=1, 2, \dots, n)$ 的并(和), 记作

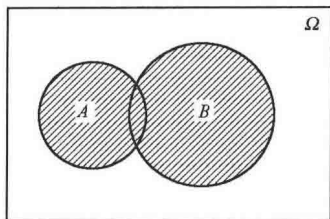


图 1—2

$$A = A_1 + A_2 + \cdots + A_n = \bigcup_{i=1}^n A_i \quad (\text{或 } A = \sum_{i=1}^n A_i).$$

事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生的事件 A 称为 $A_i (i=1, 2, \dots)$ 的并(和), 记作

$$A = A_1 + A_2 + \cdots + A_n + \cdots = \bigcup_{i=1}^{\infty} A_i \quad (\text{或 } A = \sum_{i=1}^{\infty} A_i).$$

4. 事件的交(积)

设 A, B 为两个事件, 我们把同时属于 A 及 B 的所有样本点构成的集合称为事件 A 与 B 的交或积, 记为 $A \cap B$ 或 $A \cdot B$, 有时也简记为 AB . 这就是说, 事件 $A \cap B$ 表示在一次试验中, 事件 A 与 B 同时发生. 图 1—3 中的阴影部分表示 $A \cap B$.

事件积的概念可以推广到有限个或可列无穷多个事件.

事件 $A_1, A_2, \dots, A_n$ 同时发生的事件 A 称为 $A_i (i=1, 2, \dots, n)$ 的积, 记作 $A = \bigcap_{i=1}^n A_i$ (或 $A = \prod_{i=1}^n A_i$).

事件 $A_1, A_2, \dots, A_n, \dots$ 同时发生的事件 A 称为事件 $A_i (i=1, 2, \dots)$ 的积, 记作 $A = \bigcap_{i=1}^{\infty} A_i$ (或 $A = \prod_{i=1}^{\infty} A_i$).

5. 事件的互不相容关系

设 A, B 为两个事件. 如果 $A \cdot B = \emptyset$, 那么称事件 A 与事件 B 是互不相容的(或互斥的). 这就是说, 在一次试验中事件 A 与事件 B 不能同时发生. A 与 B 互不相容的直观意义为区域 A 与区域 B 不相交, 如图 1—4 所示.

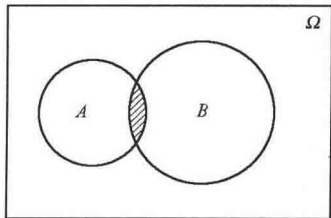


图 1—3

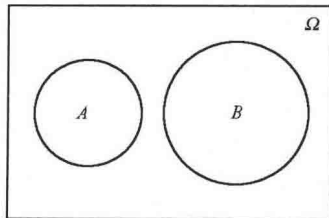


图 1—4

互不相容的概念可推广到两个以上事件: 若 $A_i A_j = \emptyset (i \neq j; i, j=1, 2, \dots)$, 称 $A_1,$

$A_2, \dots, A_n, \dots$ 互不相容. 显然, 任一随机试验中的基本事件都是互不相容的.

6. 事件的逆

对于事件 A , 我们把不包含在 A 中的所有样本点构成的集合称为事件 A 的逆 (或 A 的对立事件), 记为 $\bar{A}$. 这就是说, 事件 $\bar{A}$ 表示在一次试验中事件 A 不发生. 图 1—5 中的阴影部分表示 $\bar{A}$. 我们规定它是事件的基本运算之一.

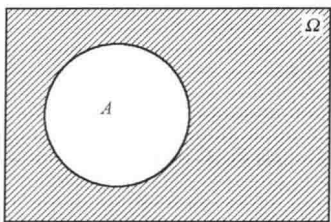


图 1—5

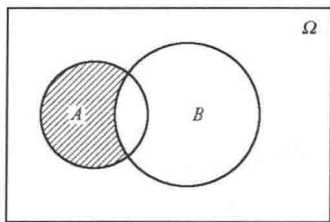


图 1—6

由于 A 也是 $\bar{A}$ 的对立事件, 因此 A 与 $\bar{A}$ 互为对立事件. 由定义可知, 两个对立事件一定是互不相容事件; 但是, 两个互不相容事件不一定是对立事件. 对立事件满足下列关系式:

$$\bar{\bar{A}} = A, \quad A\bar{A} = \emptyset, \quad A + \bar{A} = \Omega.$$

有了事件的三种基本运算我们就可以定义事件的其他一些运算. 例如, 我们称事件 $A\bar{B}$ 为事件 A 与 B 的差, 记为 $A - B$. 可见, 事件 $A - B$ 是由包含于 A 而不包含于 B 的所有样本点构成的集合. 图 1—6 中的阴影部分表示 $A - B$.

7. 完备事件组

若 n 个事件 $A_1, A_2, \dots, A_n$ 互不相容且它们的和是必然事件, 则称事件 $A_1, A_2, \dots, A_n$ 构成一个**完备事件组**. 它的实际意义是在每次试验中必然发生且仅能发生 $A_1, A_2, \dots, A_n$ 中的一个事件. 当 $n=2$ 时, A_1 与 A_2 就是对立事件. 任一随机试验的全部基本事件构成一个完备事件组.

根据上面的基本运算定义, 不难验证事件之间的运算满足以下几个规律:

(1) 交换律

$$A + B = B + A, \quad AB = BA;$$

(2) 结合律

$$A + (B + C) = (A + B) + C, \quad A(BC) = (AB)C;$$

(3) 分配律

$$A(B + C) = AB + AC, \quad A + BC = (A + B)(A + C);$$

(4) 德摩根 (De Morgan) 定理:

$$\overline{A + B} = \bar{A} \cdot \bar{B}, \quad \overline{A \cdot B} = \bar{A} + \bar{B}.$$

例 2 掷一颗骰子, 观察出现的点数. 设 A = “出现奇数点”, B = “出现点数小于 5”, C = “出现小于 5 的偶数点”. (1) 写出试验的样本空间 Ω 及事件 $A + B, A - B, AB, AC, A + \bar{C}, \bar{A} + \bar{B}$; (2) 分析事件 $A + \bar{C}, A - B, B, C$ 之间的包含、互不相容及对立关系.

解 (1) 样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$, 则

$$A = \{1, 3, 5\}, \quad B = \{1, 2, 3, 4\}, \quad C = \{2, 4\},$$

于是

$$A + B = \{1, 2, 3, 4, 5\},$$

$$A - B = \{5\},$$

$$AB = \{1, 3\},$$

$$AC = \emptyset, \quad \bar{C} = \{1, 3, 5, 6\},$$

$$A + \bar{C} = \{1, 3, 5, 6\},$$

$$\overline{A+B} = \{6\}.$$

(2) 由(1)可知, 包含关系有 $B \supset C, A + \bar{C} \supset A - B$; 互不相容的有 $A + \bar{C}$ 与 $C, A - B$ 与 C ; 事件 $A + \bar{C}$ 与 C 为对立事件.

例 3 从一批产品中每次取出一个产品进行检验(每次取出的产品不放回), 事件 A_i 表示第 i 次取到合格品($i=1, 2, 3$). 试用事件的运算符号表示下列事件: 三次都取到了合格品; 三次中至少有一次取到合格品; 三次中恰有两次取到合格品; 三次中最多有一次取到合格品.

解 三次全取到合格品: $A_1 A_2 A_3$;

三次中至少有一次取到合格品: $A_1 + A_2 + A_3$;

三次中恰有两次取到合格品: $A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3$;

三次中至多有一次取到合格品: $\bar{A}_1 \bar{A}_2 + \bar{A}_1 \bar{A}_3 + \bar{A}_2 \bar{A}_3$.

例 4 事件 A_k 表示某射手第 k 次击中目标($k=1, 2, 3$). 试用文字叙述下列事件: $A_1 + A_2$; $A_1 + A_2 + A_3$; $A_1 A_2 A_3$; $\bar{A}_2$; $A_3 - A_2$; $A_3 \bar{A}_2$; $\overline{A_1 + A_2}$; $\bar{A}_1 \bar{A}_2$; $\bar{A}_2 + \bar{A}_3$; $\overline{A_2 A_3}$; $A_1 A_2 A_3 + A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3$.

解 $A_1 + A_2$: 前两次中至少有一次击中目标;

$A_1 + A_2 + A_3$: 三次射击中至少有一次击中目标;

$A_1 A_2 A_3$: 三次射击都击中目标;

$\bar{A}_2$: 第二次射击未击中目标;

$A_3 - A_2 = A_3 \bar{A}_2$: 第三次击中目标而第二次未击中目标;

$\overline{A_1 + A_2} = \bar{A}_1 \bar{A}_2$: 前两次射击都未击中目标;

$\bar{A}_2 + \bar{A}_3 = \overline{A_2 A_3}$: 后两次射击中至少有一次未击中目标;

$A_1 A_2 A_3 + A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3$: 三次射击中至少有两次击中目标.

§ 1.2 概 率

概率论研究的是随机现象量的规律性. 因此, 仅仅知道试验中可能出现哪些事件是不够的, 还必须对事件发生的可能性的进行量的描述.

(一) 概率的统计定义

对于一般的随机事件来说,虽然在一次试验中是否发生我们不能预先知道,但是如果我们独立地重复进行这一试验就会发现不同的事件发生的可能性是有大小之分的.这种可能性的大小是事件本身固有的一种属性,这是不以人们的意志为转移的.例如,掷一枚骰子,如果骰子是匀称的,那么事件{出现偶数点}与事件{出现奇数点}的可能性是一样的;而{出现奇数点}这个事件要比事件{出现3点}的可能性更大.为了定量地描述随机事件的这种属性,我们先介绍频率的概念.

定义 1.1 在一组不变的条件 S 下,独立地重复 n 次试验 E . 如果事件 A 在 n 次试验中出现了 μ 次,则称比值 μ/n 为在 n 次试验中事件 A 出现的频率,记为 $f_n(A)$,即

$$f_n(A) = \frac{\mu}{n},$$

其中 μ 称为频数.

例如,在抛掷一枚硬币时我们规定条件组 S 为:硬币是匀称的,放在手心上,用一定的动作垂直上抛,让硬币落在一个有弹性的平面上等.当条件组 S 大量重复实现时,事件 $A=\{\text{出现正面}\}$ 发生的次数 μ 能够体现出一定的规律性.例如,进行 50 次试验出现了 24 次正面.这时

$$n = 50, \quad \mu = 24, \quad f_{50}(A) = 24/50 = 0.48.$$

一般来说,随着试验次数的增加,事件 A 出现的次数 μ 约占总试验次数的一半,换句话说事件 A 的频率接近于 $1/2$.

历史上,不少统计学家,例如皮尔逊(Pearson)等人作过成千上万次抛掷硬币的试验,其试验记录如表 1—1 所示.

表 1—1

试验者	试验次数 n	频数 m	频率 $f_n(A)$
德摩根	2 048	1 061	0.518 1
蒲 丰	4 040	2 048	0.506 9
费 勒	10 000	4 979	0.497 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5
维 尼	30 000	14 994	0.499 8

可以看出,随着试验次数的增加,事件 A 发生的频率的波动性越来越小,呈现出一种稳定状态,即频率在 0.5 这个定值附近摆动.这就是频率的稳定性,它是随机现象的一个客观规律.

可以证明,当试验次数 n 固定时,事件 A 的频率 $f_n(A)$ 具有下面几个性质:

- (1) $0 \leq f_n(A) \leq 1$;
- (2) $f_n(\Omega) = 1, f_n(\emptyset) = 0$;
- (3) 若 $AB = \emptyset$, 则

$$f_n(A+B) = f_n(A) + f_n(B).$$

定义 1.2 在相同条件下,重复进行 n 次试验,事件 A 发生的频率 $f_n(A)$ 在某个常数 p 附近摆动,而且一般来说,随着 n 的增加,这种摆动幅度会越来越小,称常数 p 为事件 A 的概率,记作 $P(A)$,即 $P(A)=p$.

如上所述,频率的稳定性是概率的经验基础,但并不是说概率决定于试验.一个事件发生的概率完全取决于事件本身的结构,是先于试验而客观存在的.

概率的统计定义仅仅指出了事件的概率是客观存在的.但我们不能用这个定义来计算事件的概率.实际上,我们可以采取一次大量试验的频率或一系列频率的平均值作为事件的概率的近似值.例如,从对一个妇产医院 6 年出生婴儿的调查中(见表 1—2)可以看到,男孩出生的频率是稳定的,可以取 0.515 作为男孩出生的概率的一个近似值.需要指出,目前国际上公认的结果为每出生 100 个女孩时,出生男孩的个数在 103~107 之间,即男孩出生率为 0.507~0.516 9.据统计,2004—2005 年我国每出生 100 个女孩,男孩数为 116.86 个,男孩出生率高达 0.538 87.这种不正常的现象是由人为因素造成的,已引起我国有关部门的高度重视.

表 1—2

出生年份	新生儿总数 n	新生儿分类数		频率(%)	
		男孩数 m_1	女孩数 m_2	男孩	女孩
1977	3 670	1 883	1 787	51.31	48.69
1978	4 250	2 177	2 073	51.22	48.78
1979	4 055	2 138	1 917	52.73	47.27
1980	5 844	2 955	2 889	50.56	49.44
1981	6 344	3 271	3 073	51.56	48.44
1982	7 231	3 722	3 509	51.47	48.53
6 年总计	31 394	16 146	15 248	51.43	48.57

虽然概率的统计定义有它的简便之处,但若试验具有破坏性,不可能进行大量重复试验时,就限制了它的应用.而对某些特殊类型的随机试验,要确定事件的概率,并不需做重复试验,而是根据人类长期积累的关于“对称性”的实际经验,提出数学模型,直接计算出来,从而给出概率相应的定义.这类试验称为**等可能概型试验**.根据其样本空间 Ω 是有限集还是无限集,可将相应的数学模型分为**古典概型**和**几何概型**.

(二) 概率的古典定义

1. 古典型随机试验

定义 1.3 若试验具有下列两个特征:

- (1) 试验的结果为有限个,即 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ (有限性);
- (2) 每个结果出现的可能性是相同的,即

$$P(\omega_i) = P(\omega_j) = \frac{1}{n} \quad (i, j = 1, 2, \dots, n) \text{ (等概性),}$$

则称此试验为**古典型随机试验**.由于这类试验曾是概率论发展初期研究的主要对象,因此

称之为古典型试验.

2. 概率的古典定义

定义 1.4 设古典型试验 E 的样本空间 Ω 有 n 个样本点, 如果事件 A 是由其中的 m 个样本点组成, 则事件 A 发生的概率 $P(A)$ 为

$$P(A) = \frac{m}{n}, \quad (1)$$

并把利用关系式(1)来讨论事件的概率的数学模型称为**古典概型**.

由古典概型的“有限性”及“等概性”两个特征, 不难看出由上述关系式给出的定义合理性. 在一次试验中, 每个样本点出现的可能性大小均为 $\frac{1}{n}$, 而事件 A 包含了 m 个样本点, 故在一次试验中, 事件 A 发生的概率应为 $m \cdot \frac{1}{n} = \frac{m}{n}$.

为了方便起见, 我们把事件 A 包含的样本点数 m 记为 m_A , 而把事件 B 包含的样本点数记为 m_B , 以示区别.

根据概率的古典定义可以计算出古典型随机试验中事件的概率. 在古典概型中确定事件 A 的概率时, 只须求出基本事件的总数 n 以及事件 A 包含的基本事件的个数 m . 为此弄清随机试验的全部基本事件是什么以及所讨论的事件 A 包含了哪些基本事件是非常重要的.

例 1 同时抛掷两枚硬币, 求下落后恰有一枚正面向上的概率.

解 设 A 表示恰有一枚正面向上的事件.

抛掷两枚硬币, 等可能的基本事件有 4 个, 即(正, 正)、(正, 反)、(反, 正)、(反, 反), 而事件 A 由其中的 2 个基本事件(正, 反)、(反, 正)组成, 故 $P(A) = \frac{1}{2}$.

例 2 设盒中有 5 只相同的玻璃杯, 其中有 3 只正品, 两只次品. 从中任取两只, 求所取出的两只都是正品的概率.

解 这里任取两只是指两只玻璃杯同时被取出, 5 只中每两只被取出的可能性相同. 若将 5 只杯子编号为 1, 2, 3, 4, 5, 前三个号代表正品, 后两个号代表次品, 于是从盒中每次任取两只的所有可能结果是

$$\begin{array}{ccccccccc} (1, 2) & (1, 3) & (1, 4) & (1, 5) & (2, 3) \\ (2, 4) & (2, 5) & (3, 4) & (3, 5) & (4, 5) \end{array}$$

其中(1, 2)表示取出 1, 2 号两只杯子, 其余类推, 因此基本事件总数 $n=10$. 设 A 表示所取出的两只都是正品的的事件, 则事件 A 含有(1, 2), (1, 3), (2, 3)三个基本事件, 即 $m=3$, 故 $P(A) = \frac{3}{10}$.

以上两例均采用列举基本事件的方法. 这种方法直观、清楚, 但较为烦琐. 在多数情况下, 由于基本事件的总数很大, 这种方法实际上是行不通的, 此时需要利用排列组合的知识解决这类问题.

例 3 从 10 件产品(其中 2 件次品, 8 件正品)之中任取 3 件, 求这 3 件产品中

- (1) 恰有 2 件次品的概率;
- (2) 至多有 1 件次品的概率.

解 设 $A = \{\text{恰有 2 件次品}\}$, $B = \{\text{至多有 1 件次品}\}$. 因为从 10 件中任取 3 件共有 C_{10}^3 种取法, 即 $n = C_{10}^3$, 而事件 A 所包含的样本点个数 $m_1 = C_2^2 C_8^1$, 事件 B 所包含的样本点个数 $m_2 = C_2^1 C_8^2 + C_2^0 C_8^3$, 所以

$$P(A) = \frac{m_1}{n} = \frac{C_2^2 C_8^1}{C_{10}^3} = \frac{1}{15}.$$

$$P(B) = \frac{m_2}{n} = \frac{C_2^1 C_8^2 + C_2^0 C_8^3}{C_{10}^3} = \frac{14}{15}.$$

例 4 从 10 件产品(其中 2 件次品, 8 件正品)中每次取 1 件观测后放回, 共取 3 次(以后简称为有放回地取 3 件). 求这 3 件产品中

- (1) 恰有 2 件次品的概率;
- (2) 至多有 1 件次品的概率.

解 设 $A = \{\text{恰有 2 件次品}\}$, $B = \{\text{至多有 1 件次品}\}$. 因为从 10 件中有放回地取 3 件共有 10^3 种取法, 而事件 A 所包含的样本点个数 $m_1 = 3 \times 2^2 \times 8$, 事件 B 所包含的样本点个数 $m_2 = 3 \times 2 \times 8^2 + 8^3$, 所以

$$P(A) = \frac{m_1}{n} = \frac{3 \times 2^2 \times 8}{10^3} = \frac{12}{125}.$$

$$P(B) = \frac{m_2}{n} = \frac{3 \times 2 \times 8^2 + 8^3}{10^3} = \frac{112}{125}.$$

例 5 设袋中有 10 个外形相同的球(其中有 6 个红球和 4 个白球), 现从中任取 3 个, 试求:

- (1) 取出的 3 个球都是红球的概率;
- (2) 取出的 3 个球中恰有一个是白球的概率.

解 本题在利用排列与组合知识计算其样本空间的样本点总数时, 相当于把外形相同的球编号, 1~6 号表示红球, 7~10 号表示白球, 把这 10 个球看成不同的球, 从中任意取 3 个球, 共有 C_{10}^3 种不同的取法, 每种取法都对应一个样本点, 所以, 该试验的样本点总数为 $n = C_{10}^3$.

(1) 设 $A = \{\text{取出的 3 个球都是红球}\}$, 而事件 A 包含了 $m_A = C_6^3$ 个样本点, 则

$$P(A) = \frac{m}{n} = \frac{C_6^3}{C_{10}^3} = \frac{1}{6}.$$

(2) 设 $B = \{\text{取出的 3 个球中恰有一个是白球}\}$, 而事件 B 包含的样本点数 $m_B = C_4^1 C_6^2$, 则

$$P(B) = \frac{m}{n} = \frac{C_4^1 C_6^2}{C_{10}^3} = \frac{1}{2}.$$

本题称为随机取球问题, 古典概型大部分问题都能用随机取球“模型”来描述. 例如, 把例 3、例 4 中的产品看成球, 则产品抽样检查就是其中之一.

例 6 设袋中有外形相同的 a 个白球和 b 个红球, 现从中每次取 1 个看后不放回(以