普通高等教育“十三五”规划教材

经济数学基础教程

概率论与数理统计

张从军　刘亦农
肖丽华　周惠新　编

科学出版社

普通高等教育“十三五”规划教材

经济数学基础教程

概率论与数理统计

张从军　刘亦农
肖丽华　周惠新　编

科 学 出 版 社

北 京

内 容 简 介

本书是“经济数学基础教程”之一. 主要内容包括随机事件与随机变量、二维随机变量及其联合概率分布、随机变量的数字特征、统计估计方法、统计检验方法、一元线性回归分析与方差分析等各章，并配有适量习题.

本书贯彻问题教学法的基本思想，对许多数学概念，先从提出经济问题入手，再引入数学概念，介绍数学工具，最后解决所提出的问题，从而使学生了解应用背景，提高学习的积极性；书中详细介绍相应的数学软件，为学生将来的研究工作和就业奠定基础；穿插于全书的数学建模的基本思想和方法，引导学生学以致用，学用结合.

本书可作为普通高等学校财经类各专业概率论与数理统计课程的教材，最大限度地适应财经类各专业学习该课程和后续课程的需要，以及报考研究生的需要和将来从事财经有关实际工作的需要.

图书在版编目(CIP)数据

概率论与数理统计/张从军等编. —北京：科学出版社，2016.8
普通高等教育“十三五”规划教材·经济数学基础教程
ISBN 978-7-03-049268-5

Ⅰ. ①经… Ⅱ. ①张… Ⅲ. ①经济数学—高等学校—教材 ②概率论—高等学校—教材 ③数理统计—高等学校—教材 Ⅳ. ①F224 ②O21

中国版本图书馆 CIP 数据核字(2016) 第 144286 号

责任编辑：姚莉丽 张中兴／责任校对：邹慧卿
责任印制：白 洋／封面设计：陈 敬

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

大厂书文印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2016 年 8 月第 一 版 开本：720 × 1000 B5
2016 年 8 月第一次印刷 印张：22 1/2
字数：451 000

定价：39.00 元

(如有印装质量问题，我社负责调换)

前言

随着社会经济的迅猛发展, 随着数学在经济活动和经济研究中的作用日益凸显, 随着数学的理论和方法越来越广泛地应用到自然科学、社会科学和工程技术的各个领域, 对高等学校财经类各专业培养人才的数学素养要求越来越高. 经济数学基础课程, 在提高财经类专业人才的数学素养方面, 起着至关重要的基础性作用. 这类课程的思想和方法, 是人类文明发展史上理性智慧的结晶, 它不仅提供的是解决实际问题的有力数学工具, 同时还提供给学生一种思维的训练, 帮助学生具备作为复合型、创造型、应用型人才所必须的文化素质和修养.

怎样使经济数学基础课程充分发挥上述作用, 怎样使经济数学基础课程更趋符合培养目标的课程体系, 怎样兼顾经济数学基础课程的理论性与应用性、思想性与工具性, 怎样突出经济数学基础课程的财经类专业特色, 现有的经济数学基础教材固然很多, 但要处理好以上问题, 仍需认认真真地思考与探索, 仍有大量的工作要做. 作为我们主持承担的全国高等教育科学“十五”规划重点研究课题的研究内容之一, 我们从 2002 年始, 着手陆续编写一套经济数学基础课程和经济数学应用课程教材, 本书是其中的概率论与数理统计部分. 该教材于 2007 年 1 月在复旦大学出版社出版了第一版.

我们在编写思想, 体系安排, 内容取舍, 教学方法诸方面按照上述要求作了一些尝试. 特别注重了以下几点:

(1) 最大限度的适应财经类各专业学习该课程的需要, 后续课程的需要, 报考研究生的需要, 将来从事与财经有关的实际工作的需要.

(2) 贯彻问题教学法的基本思想, 对许多数学概念, 先从提出经济问题入手, 再引入数学概念, 介绍数学工具, 最后解决所提出的问题. 使学生了解应用背景, 提高学习的积极性.

(3) 详细介绍相应的数学软件, 为学生将来的研究工作和就业奠定基础.

(4) 穿插了数学建模的基本思想和方法, 引导学生学以致用, 学用结合; 编写了附录部分, 增加了数学文化的内容.

本书由张从军教授提出编写思想和编写提纲、列出章节目录、编写附录部分, 最后对全书进行修改补充、统稿、定稿. 肖丽华副教授编写了第 1、2 章, 刘亦农副教授编写了第 3、6 章以及全书的软件部分内容, 周惠新副教授编写了第 4、5 章.

该教材自 2007 年出版以来, 得到了许多院系、教师和广大学生的充分肯定. 经过多年使用, 我们陆续收到了许多读者特别是一些一线任课教师的宝贵意见, 同时

我们也发现了不少需要修改与提高之处.

我们一直认为, 编写一本教材似乎不难, 但编写一本适用的教材绝非易事. 编写此类经济数学基础教材和经济数学应用教材更不是一劳永逸、一蹴而就的事. 因此, 既要保持相对的连续性和稳定性, 又要紧紧围绕人才培养的目标体系、课程体系, 吸收最新的有关教学科研成果, 不断修改完善. 为了满足广大学生的实际使用需要, 更好地兼顾教材内容的思想性与工具性、科学性与可接受性、先进性与适用性, 更有利于提高学生的数学素养和应用能力, 我们需要对教材内容作进一步的精雕细琢.

作为我们主持承担的教育部高等理工教育数学教学研究与改革课题 (教高司 [2007]143 号)、江苏省高等教育教改立项研究课题 (苏教高 [2007]18 号) 的研究内容之一, 适应逐步打造精品教材的要求, 我们于 2011 年 11 月在复旦大学出版社出版了该教材第二版. 通过再版, 我们改进了有关内容的表述, 调整了某些结构顺序, 充实了一定例题和习题, 特别是进一步体现了经管专业使用特色.

作为江苏省重点教材的系列教材, 根据江苏省重点教材的要求, 我们将修定后的教材交由科学出版社于 2016 年纳入普通高等教育“十三五”规划教材新版.

复旦大学数学学院博导应坚刚教授、管理学院统计系主任郑明教授审阅了该书第一版书稿并提出了宝贵意见, 南京财经大学经济学院陈耀辉教授阅读了书稿并提出了有益的建议、南京财经大学应用数学学院的一些经济数学基础课程任课教师在教学中不断完善了本书内容. 复旦大学出版社范仁梅总监等不辞劳苦, 数次往返于上海与南京之间, 对该教材的出版给予了大力支持. 编者在此向他们表示衷心感谢!

值此该教材新版之际, 我们还要特别感谢科学出版社对该教材出版给予的大力支持. 感谢使用该教材的教师和读者给我们提出的宝贵意见, 感谢相关院系对我们的支持和帮助, 感谢关心该教材建设的有关校领导和教务部门, 感谢审定该教材的相关专家.

本书在编写过程中, 参考了大量的相关教材和资料, 选用了其中的有关内容和例题、习题, 在此谨向有关编者、作者一并表示我们的谢意.

最后, 我们再次恳切期望有关专家、学者不吝赐教, 诚恳期望使用该教材的教师和同学们, 提出并反馈宝贵意见.

联系邮箱: yysxx@njue.edu.cn.

编　者

2016 年 4 月

目录

第 1 章　随机事件与随机变量

> 概率论是"生活真正的领路人"，如果没有对概率的某种估计, 那么我们就寸步难移, 无所作为.
>
> —— 杰文斯 (W.S.Jevons)

在自然界和人类的社会生活与生产实践中存在着大量的随机现象, 虽然这些现象具有偶然性但因其存在规律性, 使人们对它们的研究发生了兴趣, 概率论与数理统计就是一门以随机现象及其规律性为研究对象的数学学科. 人们希望以它为理论依据对现实生活中的某些事物进行统计推断从而做出正确的决策.

本章从实际问题出发, 介绍概率论中两个最基本的概念: 随机事件及其概率; 进而讨论两类随机变量及其概率分布; 最后举例介绍所涉上述问题的软件应用. 本章内容是学习概率论的基础.

1.1　从推断问题和信用卡管理谈起

1.1.1　推断问题

某市政府信访办公室承诺在国庆节放假期间仍然安排值班人员接待来访群众. 记者发现在 10 月 1 日至 10 月 7 日的七天内信访办接待了 12 名来访者, 记录显示他们是在 10 月 2 日和 10 月 4 日两天来访的, 记者的疑问是这七天长假期间是否每天都有工作人员在值班.

假设每位来访者可选择 1 日至 7 日的任何一天来访, 则 12 名来访者共有 7^{12} 种组合方式来到信访办公室, 而他们均在 2 日和 4 日来访的组合方式共有 2^{12} 种, 由此可知 12 名来访者都在这两天来访的可能性为 $\left(\dfrac{2}{7}\right)^{12}$, 约为 0.0000003. 这么小的可能性可推断七天长假并不是每天均有人值班, 记者的怀疑是有道理的.

1.1.2　信用卡管理问题

信用卡发行是银行重要的业务之一, 一方面银行希望争取尽量多的客户, 另一方面却是信用卡客户透支问题, 从而信用卡管理是一个重要问题.

某银行将客户分为 (信用) 好和 (信用) 坏两类, 并通过分析历史数据得到: 在每个月都会有近 1%的好客户和 10%的坏客户透支银行账户. 当一位新客户来银

行开办现金账户时, 信用处通过基本检验后认为这位客户大概有 70%的机会是一位好客户. 问题是这位客户在第一个月内就透支, 请问银行对这位客户的信用度有什么改变? 如果这位客户在第二个月仍透支呢?

假设 $H=$ “信用好”, $T=$ “透支其账户”, 银行的历史数据显示:

$$P(T\mid H)=0.01,\quad P\left(T\mid\overline{H}\right)=0.1.$$

另一方面, 银行关于这位客户最初的信用观点是

$$P(H)=0.7.$$

由贝叶斯 (Bayes) 理论, 有

$$P(H\mid T)=\frac{P(H)\cdot P(T\mid H)}{P(H)\cdot P(T\mid H)+P\left(\overline{H}\right)\cdot P\left(T\mid\overline{H}\right)}=\frac{7}{37}\approx 18.92\%.$$

银行认为他是好客户的可能性由 70%降到不足 20%. 在这里 $P(H)=0.7$ 称为先验概率, $P(H\mid T)=\dfrac{7}{37}$ 称为后验概率. 下面我们来考虑第二个月, 在第二个月内这位客户透支, 此时

$$P(H)=\frac{7}{37},$$

$$\begin{aligned}P(H\mid T)&=\frac{P(H)\,P(T\mid H)}{P(H)\,P(T\mid H)+P\left(\overline{H}\right)P\left(T\mid\overline{H}\right)}\\&=\frac{\dfrac{7}{37}\times 0.01}{\dfrac{7}{37}\times 0.01+\dfrac{30}{37}\times 0.1}=\frac{7}{307}\approx 2.28\%.\end{aligned}$$

此时银行不会再认为他是一位 (信用) 好的客户了.

通过上述两个问题可以看到, 人们对现实世界的种种认识很多情况下是对各种事件发生可能性大小的判断. 反过来, 这些事件发生的可能性大小又影响着人们的行为, 我们的生活离不开概率. 以下从最基本的内容开始讨论.

1.2　随机事件及其概率

1.2.1　样本空间

在自然界和人类社会活动中存在着许多现象, 其中有些现象只要满足一定的条件就必然发生. 例如: “在标准大气压下, 纯水加热到 100°C 时会沸腾”, “在没有外力作用的条件下静止的物体必然静止”. 这类现象称为**确定性现象**.

自然界和人类社会活动中还广泛存在着与确定性现象有着本质区别的另一类现象, 例如: 掷一枚硬币, 可能正面朝上也可能反面朝上; 某城市明天发生交通事故的次数; 从生产线下来的产品是否为合格品, 这种在同样条件下进行同样的观测或实验却可能发生不同结果的现象称为**随机现象**. 这种普遍存在的看起来好像毫无规律的随机现象后面实际却隐藏着某种规律性. 例如, 多次重复抛一枚硬币得到正面朝上大致有一半, 某城市明天发生交通事故的次数按照一定规律分布等等. 这种在大量重复试验或观察中所呈现出的固有规律性称为**统计规律性**. 概率论和数理统计就是研究随机现象统计规律性的一门学科.

一般地, 使随机现象得以实现及对它观察的全过程通称为**随机试验**, 简称**试验**, 记为 E. 要完成一个随机试验, 主要是明确实现它的"一定条件"以及由它产生的一切可能的"基本结果". 这里的"一定条件"可以是人为的也可以是客观存在的; 这里的"基本结果"是指随机实验最简单的, 不可 (或不必) 再细分的结果.

定义 1.1 随机试验的每一个基本结果称为**样本点**, 记作 $\omega_1, \omega_2, \cdots$ 随机试验的所有样本点组成的集合称为**样本空间**, 记作 Ω, 即

$$\Omega = \{\omega_1, \omega_2, \cdots\}.$$

例 1.1 E: "掷一枚硬币观察其朝上的面"; 可能出现的结果是正面或反面; $\Omega = \{\text{正}, \text{反}\}$.

例 1.2 E : "一个人进行射击, 记录他直至击中目标的射击次数"; 可能结果是 $1, 2, \cdots$; $\Omega = \{1, 2, \cdots\}$.

例 1.3 E : "观察一只灯泡的使用寿命"; 可能出现的结果是任一非负正数; $\Omega = [0, +\infty)$.

例 1.4 E : "观测某市每日的最高气温和最低气温"; 以 x, y 分别表示最高和最低气温, 人们总可以确定此地气温的上界 a 和下界 b, 可能出现的结果是坐标平面中的一个三角形; $\Omega = \{(x, y) | b \leqslant y \leqslant x \leqslant a\}$.

不难看出, 样本空间 Ω 可以是数集, 也可以是任何抽象的集合; 可以是有限集, 可列集, 也可以是不可列的无穷集合; 可以是一维的也可以是多维的集合. 所以, 正确地确定不同随机试验的样本点和样本空间是非常重要的.

1.2.2 随机事件及其运算

在一次随机试验中, 我们通常关心的是带有某些特征的那些样本点所组成的集合. 例如: 例 1.2 中"3 次以内击中目标"; 例 1.3 中的"一只灯泡的寿命超过 500 小时". 这种带有某种特征的样本点组成的集合称为**随机事件**, 通常用大写字母 $A, B, C, \cdots$ 等表示. 因此, 可记

$$A = \{\text{一个人射击, 3 次以内击中目标}\} = \{1, 2, 3\},$$

$$B=\{\text{一只灯泡的寿命超过 500 小时}\}=(500,+\infty).$$

随机事件是样本空间 Ω 的一个子集, 在试验中, 如果事件 A 包含的某一个样本点 ω 出现了, 则称 A 发生, 记为 $\omega\in A$. 由样本空间 Ω 中的单个元素组成的子集称为**基本事件**, 而样本空间 Ω 的最大子集 Ω 称为**必然事件**, 样本空间 Ω 的最小子集空集 $\varnothing$ 称为**不可能事件**.

一个样本空间 Ω 中, 可以有很多随机事件, 人们通常研究这些事件的关系及运算, 以便通过较简单的事件的统计规律去研究较复杂事件的统计规律. 下面介绍事件间的关系及运算, 它们与集合论中集合之间的关系及运算是一致的.

(1) 包含

如果事件 A 发生必然导致事件 B 发生, 则称事件 B **包含**事件 A, 记作 $B\supset A$ 或 $A\subset B$. 对任一事件 A, 有 $\varnothing\subset A\subset\Omega$.

如果 $A\subset B$ 且 $B\subset A$, 则称事件 A 与事件 B **相等**, 记作 $A=B$.

(2) 和

两个事件 A 与 B 中至少有一个发生, 称为事件 A 与事件 B 的**和**, 记作 $A+B$ 或 $A\cup B$.

$\sum\limits_{i=1}^{n}A_i=A_1+A_2+\cdots+A_n$ 表示 n 个事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生;

$\sum\limits_{i=1}^{\infty}A_i=A_1+A_2+\cdots+A_n+\cdots$ 表示可列个事件$A_1,A_2,\cdots,A_n,\cdots$ 中至少有一个发生.

(3) 积

两个事件 A 与 B 同时发生, 称为事件 A 与 B 的**积**, 记作 AB 或 $A\cap B$.

$\prod\limits_{i=1}^{n}A_i=A_1A_2\cdots A_n$ 表示 n 个事件 $A_1,A_2,\cdots,A_n$ 同时发生;

$\prod\limits_{i=1}^{\infty}A_i=A_1A_2\cdots A_n\cdots$ 表示可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 同时发生.

(4) 互不相容

如果事件 A 与 B 不能同时发生, 则称事件 A 与 B **互不相容或互斥**, 此时必有 $AB=\varnothing$. 基本事件是互不相容的.

如果 n 个事件 $A_1,A_2,\cdots,A_n$ 中任意两个事件都互不相容, 即

$$A_iA_j=\varnothing\quad(1\leqslant i<j\leqslant n),$$

则称这 n 个事件是互不相容的或互斥的.

(5) 对立 (逆)

如果两个事件 A 与 B 满足 $A+B=\Omega$, $AB=\varnothing$, 则称事件 A 是事件 B 的**对立事件或逆事件**. 此时事件 B 也是事件 A 的逆事件, 所以事件 A 与事件 B 是互逆事件, 记作 $\overline{A}=B$ 或 $\overline{B}=A$. 显然 $\overline{\Omega}=\varnothing$, $\overline{\varnothing}=\Omega$, $\overline{\overline{A}}=A$.

(6) 差

如果事件 A 发生且事件 B 不发生, 称为事件 A 与 B 的**差**, 记为 $A-B$. 显然 $A-B=A-AB=A\overline{B}$.

维恩 (John Venn) 图 (图 1-1) 能更直观地将事件间的关系及运算表示出来. 用一矩形表示样本空间 Ω, 矩形内的点表示样本点, 矩形区域内的子集 A,B 代表随机事件 A,B.

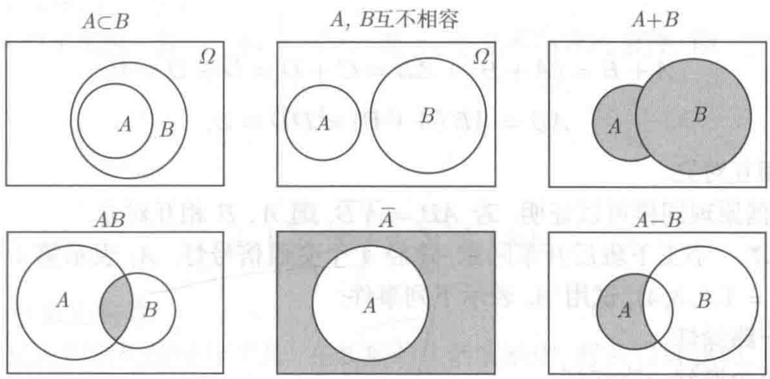

图 1-1

对于事件的运算有如下的运算规律.

事件的运算律

(1) **交换律**: $A+B=B+A$, $AB=BA$;

(2) **结合律**: $(A+B)+C=A+(B+C)$, $(AB)C=A(BC)$;

(3) **分配律**: $(A+B)C=AC+BC$;

(4) **对偶原理**: $\overline{A+B}=\overline{A}\,\overline{B}$, $\overline{AB}=\overline{A}+\overline{B}$.

对于 n 个事件, 甚至对于可列个事件, 对偶原理也成立.

例 1.5 证明 $\overline{A+B}=\overline{A}\,\overline{B}$.

证法一

$$\begin{aligned}\overline{A+B}&=\overline{\{A,B\text{至少有一个发生}\}}=\{A,B\text{都不发生}\}\\&=\{\overline{A},\overline{B}\text{都发生}\}=\overline{A}\,\overline{B}.\end{aligned}$$

证法二 $\forall\omega\in\overline{A+B}$, 有 $\omega\in\Omega-(A+B)$, 即

$$\omega\in\Omega-A\quad\text{且}\quad\omega\in\Omega-B,$$

所以 $\omega \in \overline{A}$ 且 $\omega \in \overline{B}$, 即 $\omega \in \overline{A}\,\overline{B}$, 亦即 $\overline{A+B} \subset \overline{A}\,\overline{B}$.

反之, $\forall \omega \in \overline{A}\,\overline{B}$, 有 $\omega \in \overline{A}$ 且 $\omega \in \overline{B}$, 即

$$\omega \in \Omega - A \quad 且 \quad \omega \in \Omega - B,$$

于是 $\omega \in \Omega - A - B$, 即 $\omega \in \overline{A+B}$, 所以 $\overline{A}\,\overline{B} \subset \overline{A+B}$.

因此 $\overline{A+B} = \overline{A}\,\overline{B}$.

例 1.6　设 A, B 是随机事件, 若满足 $A+B = \overline{A} + \overline{B}$, 证明: A, B 相互对立.

证明　我们只要证明 $A+B=\Omega$ 且 $AB=\varnothing$ 即可.

记 $C=A+B, D=AB$, 则由等式

$$\overline{AB} = \overline{A} + \overline{B} = A + B$$

可得

$$A+B=(A+B)+AB=C+D=\overline{D}+D=\Omega,$$
$$AB=AB\,(A+B)=D\overline{D}=\varnothing,$$

即 A, B 相互对立.

由对偶原理同样可以证明: 若 $AB=\overline{A}\,\overline{B}$, 则 A、B 相互对立.

例 1.7　小王下班后开车回家, 途经 4 个交通信号灯. A_i 表示第 i 个路口遇上红灯 $(i=1,2,3,4)$, 试用 A_i 表示下列事件:

(1) 一路绿灯;

(2) 至少遇到一次红灯;

(3) 只遇到一次红灯;

(4) 至少遇到 3 次红灯;

(5) 至多遇到 3 次红灯.

解　(1) $\overline{A_1}\,\overline{A_2}\,\overline{A_3}\,\overline{A_4} = \overline{A_1+A_2+A_3+A_4}$.

(2) $A_1+A_2+A_3+A_4$.

(3) $A_1\overline{A_2}\,\overline{A_3}\,\overline{A_4} + \overline{A_1}A_2\overline{A_3}\,\overline{A_4} + \overline{A_1}\,\overline{A_2}A_3\overline{A_4} + \overline{A_1}\,\overline{A_2}\,\overline{A_3}A_4$.

(4) $\overline{A_1}A_2A_3A_4 + A_1\overline{A_2}A_3A_4 + A_1A_2\overline{A_3}A_4 + A_1A_2A_3\overline{A_4} + A_1A_2A_3A_4$.

(5) $\overline{A_1A_2A_3A_4} = \overline{A_1} + \overline{A_2} + \overline{A_3} + \overline{A_4}$.

例 1.8　掷一枚骰子 (6 个面), 观察其朝上面的数字. 设事件 $A=\{$出现奇数点$\}$, 事件$B=\{$出现偶数点$\}$, $C=\{$小于 4 点$\}$, 求: 事件 $A+C$, BC, $A-C$, AB, $A+B$, 并问 A, B, C 中哪两个事件是对立的?

解　因为 $\Omega=\{1,2,3,4,5,6\}$, $A=\{1,3,5\}$, $B=\{2,4,6\}$, $C=\{1,2,3\}$, 所以

$$A+C=\{1,2,3,5\}, \quad BC=\{2\},$$
$$A-C=\{5\}, \quad AB=\varnothing, \quad A+B=\Omega.$$

因此, A 与 B 是对立事件.

1.2.3 频率与概率

定义 1.2 设 E 为任一随机试验, A 为 E 中的一随机事件. 在 n 次重复试验中 A 出现的频数 (次数) 记为 $\mu_n(A)$, 称比值

$$f_n(A)=\frac{\mu_n(A)}{n} \tag{1.1}$$

为随机事件 A 在 n 次试验中出现的**频率**.

易知频率 $f_n(A)$ 具有下述性质.

频率 $f_n(A)$ 的性质

(1) 非负性: $f_n(A)\geqslant 0$;

(2) 规范性: $f_n(\Omega)=1$;

(3) 有限可加性: 若 $A_1,A_2,\cdots,A_m$ 是 m 个互不相容的事件, 即

$$A_iA_j=\varnothing,\quad i\neq j,$$

则有

$$f_n\left(\sum_{i=1}^{m}A_i\right)=\sum_{i=1}^{m}f_n(A_i),$$

其中 n 为任意正整数.

频率是人们常用的统计工具. 在长期的生活实践中, 经验告诉我们: 事件 A 发生的可能性越大, 其发生的频率也越大; 反之, 事件 A 发生的频率越大, 可以想象事件 A 发生的可能性也越大. 虽然频率的数值不固定, 但当试验次数 n 充分大时, 事件 A 发生的频率将在某个实数的附近波动, 一般来说, 试验次数越多, 波动越小, 这称作频率的稳定性.

在掷一枚硬币时, 既可能出现正面也可能出现反面, 假如硬币是均匀的, 直观感觉是出现正面和出现反面的可能性相等, 即在大量的试验中出现正面的频率应接近 50%, 为了验证这种频率的稳定性, 历史上有不少统计学家做过试验, 试验结果如表 1-1 所示.

表 1-1

实验者	掷硬币次数	出现正面次数	频率
蒲丰 (Buffon)	4040	2048	0.5069
皮尔逊 (Pearson)	12000	6019	0.5016
皮尔逊 (Pearson)	24000	12012	0.5005

定义 1.3 在重复进行同一试验时, 随着试验次数 n 的无限增大, 事件 A 发生的频率 $f_n(A)=\dfrac{\mu_n(A)}{n}$ 会稳定在某一实数值附近. 这个实数值反映了事件 A 发

生的可能性的大小, 称为事件 A 的概率, 记作 $P(A)$.

这种通过频率的稳定性得到的概率定义称为概率的统计定义. 它说明频率与概率有着密切的联系. 概率是事件的一种客观存在而频率是概率的近似值.

1.2.4 古典概型

有一类问题, 人们可以根据问题本身所具有的某种"对称性"来分析问题的本质, 就可以计算出其概率了. 例如, 掷一枚硬币, 可能出现正面也可能出现反面, 设 $P(A)$ 表示出现正面的概率, $P(\overline{A})$ 表示出现反面的概率, 若硬币是均匀的, 那么两者发生的可能性应该均等, 所以 $P(A)=P(\overline{A})=\frac{1}{2}$. 这类问题便是最简单的古典概型的问题.

定义 1.4 如果一随机试验具有如下特点:

(1) 样本空间只含有有限个样本点, 即 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$;

(2) 各样本点发生的可能性相等, 即有

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)=\frac{1}{n},$$

则称此随机试验为**古典概型**. 对任一随机事件 A 来说, 如果 A 包含了 m 个样本点, 则 A 发生的概率

$$P(A)=\frac{m}{n}=\frac{A\text{ 包含的样本点数}}{\Omega\text{ 中的样本点总数}}. \tag{1.2}$$

(1.2) 式称为**概率的古典定义**.

定理 1.1 古典概率有以下性质:

(1) 非负性 $P(A)\geqslant 0$; (1.3)

(2) 规范性 $P(\Omega)=1$; (1.4)

(3) 若 A 与 B 互不相容, 即 $AB=\varnothing$, 则

$$P(A+B)=P(A)+P(B). \tag{1.5}$$

证明 (1), (2) 显然, 仅证明 (3).

设 A 包含有 m_1 个样本点, B 包含有 m_2 个样本点, 因为 $AB=\varnothing$, 所以 $A+B$ 中包含有 m_1+m_2 个样本点. 故

$$P(A+B)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}=P(A)+P(B).$$

推论 (1) $P(\overline{A})=1-P(A)$; (1.6)

(2) 有限可加性: 设事件 $A_1,A_2,\cdots,A_m$ 两两互不相容, 即 $A_iA_j=\varnothing, i\neq j$, 则

$$P\left(\sum_{i=1}^{m}A_i\right)=\sum_{i=1}^{m}P(A_i). \tag{1.7}$$

例 1.9 将一枚硬币抛掷三次. 设 A_1 为"恰有 2 次正面", A_2 为"至少有一次正面", 求 $P(A_1)$ 和 $P(A_2)$.

解 样本空间 $\Omega=$\{(正, 正, 正), (正, 正, 反), (正, 反, 正), (反, 正, 正), (正, 反, 反), (反, 正, 反), (反, 反, 正), (反, 反, 反)\}.

样本空间中包含有限个元素, 且由对称性知每个基本事件发生的可能性相同, 故由 (1.2) 式知

$$P(A_1)=\frac{3}{8};\quad P(A_2)=\frac{3+3+1}{8}=\frac{7}{8},$$

或

$$P(A_2)=1-P\left(\overline{A_2}\right)=1-\frac{1}{8}=\frac{7}{8}.$$

如果上例中样本空间设为 $\Omega=$\{0 个正面, 1 个正面, 2 个正面, 3 个正面\}, 那么就不满足每个基本事件发生的可能性相同的古典概型要求了. 因此要计算古典概率中随机事件的概率, 首先要确定试验中可能出现的基本事件总数 n 即样本空间的构成, 其次要弄清楚所求随机事件 A 所含有的基本事件数 m, 再用 (1.2) 式得到随机事件 A 的概率.

例 1.10 某商场周末促销设立了刮刮奖. 每位在商场一次消费满 100 元以上的顾客可以获得一张刮刮卡, 号码是由 0~9 这 10 个数字组成的 4 位数, 吉祥数字是 6, 8, 9. 一等奖是 4 个相同的吉祥数; 二等奖是 4 个吉祥数中 3 个相同的; 三等奖是 4 个吉祥数中 2 个相同的. 请问一次刮奖分别刮出一、二、三等奖的概率.

解 设 $A_i=\{$中第 i 等奖$\}$, $i=1,2,3$.

由于每一位数字有 10 种选择, 因此样本总数为 10^4. 当然 4 个相同的吉祥数字只有 6666, 8888, 9999 这 3 种情况, 则有

$$P(A_1)=\frac{3}{10^4}=0.0003;$$

二等奖是 3 个吉祥数中有 1 个出现 3 次, 另外 2 个出现 1 次的选排列, 则

$$P(A_2)=\frac{\mathrm{C}_4^3\cdot\mathrm{A}_3^2}{10^4}=0.0024;$$

三等奖是 3 个吉祥数中有 2 个相同, 剩下 2 个可以相同也可以不相同, 则

$$P(A_3)=\frac{\mathrm{C}_4^2\mathrm{C}_3^2+\mathrm{C}_4^2\mathrm{C}_3^1\mathrm{A}_2^2}{10^4}=0.0054.$$

这种活动的中奖面近 1%. (在计算 4 个数是两两相同的情形时需注意在 4 个格子中选 2 个格子放一个数, 剩下的 2 个格子放一个数, 例如, 6868, 选中第 1、第 3 格子放 6, 剩余放 8 与选中第 2、第 4 格子放 8 剩余放 6 是相同情况, 不能重复计算. 把选中格子与剩余格子看作 2 个整体. 放入 2 个吉祥数字, 有 C_3^2 种情况.)

例 1.11　设有一批产品共 100 件, 其中有 5 件次品, 现从中任取 5 件, 问:

(1) 有 2 件次品的概率是多少?

(2) 至少有 1 件次品的概率是多少?

解　设 $A=\{5$ 件产品中有 2 件次品$\}$, $B=\{5$ 件产品中至少有 1 件次品$\}$. 从 100 件产品中任取 5 件共有 C_{100}^5 种不同的结果, 总样本数为 C_{100}^5.

(1) 5 件产品中有 2 件次品 3 件正品, 则

$$P(A)=\frac{\mathrm{C}_5^2\cdot\mathrm{C}_{95}^3}{\mathrm{C}_{100}^5}\approx 1.8\%.$$

(2) 5 件产品中至少有 1 件次品包括多种情况, 不方便计算, 可以利用它的逆事件即没有 1 件次品计算, 则

$$P(B)=1-P(\overline{B})=1-\frac{\mathrm{C}_{95}^5}{\mathrm{C}_{100}^5}\approx 23.0\%.$$

例 1.12 (抽样检验)　如果某批产品中有 a 件次品和 b 件正品, 我们采用有放回和不放回抽样方式从中抽 n 件产品, 问正好有 $m(m\leqslant a)$ 件次品的概率各是多少?

解　所求概率与抽样方式有关, 下面分别讨论.

在**有放回抽样**中, 被抽中的产品检验后仍放回产品堆中, 再抽第二次, 因此这件产品以后仍能再次被抽到, 样本点是可重复的排列, 共有 $(a+b)^n$, 次品恰好出现 m 次共有 $\mathrm{C}_n^m a^m b^{n-m}$ 种情况, 故

$$P(\text{恰有 } m \text{ 个次品})=\frac{\mathrm{C}_n^m a^m b^{n-m}}{(a+b)^n}=\mathrm{C}_n^m\left(\frac{a}{a+b}\right)^m\left(\frac{b}{a+b}\right)^{n-m}.$$

在**不放回抽样**中, 被抽中的产品不再放回, 因而以后不会再被抽到, 从 $a+b$ 个产品中取出 n 个产品的组合共 C_{a+b}^n 种情况, n 件产品中恰有 m 件是次品则剩下的 $n-m$ 件应是正品, 故

$$P(\text{恰有}m\text{个次品})=\frac{\mathrm{C}_a^m\cdot\mathrm{C}_b^{n-m}}{\mathrm{C}_{a+b}^n}.$$

在实际工作中, 当产品总数很大而抽样数不大时, 往往采用不放回方式抽样, 而用有放回方式计算, 此时两者数值相差不大.

例 1.13 (分房问题)　设有 n 个人, 每个人都等可能被分配到 N 个房间中的任意一间去住 $(n\leqslant N)$, 求下列事件的概率:

(1) 指定的 n 个房间各有一人住;

(2) 恰有 n 个房间, 每间各住一人;

(3) 某个指定的房间有 $k(k\leqslant n)$ 人住.