

C21-43
X47

全国高等院校数学基础课教材

新编概率论与数理统计

主编 肖筱南

编著 茹世才 欧阳克智

王惠君 肖筱南



A0963495

北京大学出版社

· 北 京 ·

图书在版编目(CIP)数据

新编概率论与数理统计/肖筱南主编. —北京: 北京大学出版社, 2002. 1

ISBN 7-301-05263-4

I. 新… II. 肖… III. ① 概率论-高等学校-教材 ② 数理统计-高等学校-教材 IV. 021

中国版本图书馆 CIP 数据核字(2001)第 072294 号

书 名: 新编概率论与数理统计

著作责任者: 肖筱南 主编

责任编辑: 刘 勇

标准书号: ISBN 7-301-05263-4/O · 0517

出版者: 北京大学出版社

地 址: 北京市海淀区中关村北京大学校内 100871

网 址: <http://cbs.pku.edu.cn/cbs.htm>

电 话: 出版部 62752015 发行部 62754140 邮购部 62752019

电子信箱: zpup@pup.pku.edu.cn

印刷者: 北京大学印刷厂

发 行 者: 北京大学出版社

经 销 者: 新华书店

850×1168 32 开本 13.125 印张 320 千字

2002 年 1 月第 1 版 2002 年 1 月第 1 次印刷

印 数: 0001—11000 册

定 价: 19.00 元

前 言

概率论与数理统计是从数量方面研究随机现象规律性的数学学科.随着知识经济、信息时代的来临,研究随机现象的数学理论、方法及其应用愈来愈广泛,概率统计的思想方法已经渗透到了自然科学与社会科学的各个领域,并且还在继续拓广.

现代科学技术的迅速发展,使知识更新的速度不断加快.对此,作为高等工科、经济类院校重要基础课的概率统计教材应该如何适应当代知识经济发展的需要,不断开拓创新,已成为摆在我们面前的紧要任务.为了进一步提高概率统计的教学质量,更好地适应新世纪对高等工科及经济类院校培养复合型高素质人才的需要,我们在概率统计课程的教材建设与改革中,结合多年的教学研究与实践,在博采众家之长的基础上,紧扣教育部最新颁布的教学大纲,针对工科类、经济类院校学生学习概率统计的实际需要,几经修改编写了本书.在编写过程中,考虑到工科与经济类数学教学的特点以及学生报考工学、经济学硕士研究生入学考试的需要,我们在遵循本学科系统性与科学性的前提下,尽量注意贯彻由浅入深、循序渐进、融会贯通的教学原则与直观形象的教学方法,既注重概率统计基本概念、基本理论和方法的阐述,又注重学生基本运算能力的训练和分析问题、解决问题能力的培养,以达到既便于教师教学,又便于学生自学之目的.

为了帮助各类学生更好地掌握本课程内容,本书每节分层次配有(A)基本练习题,(B)带有一定难度的提高练习题,此外,每章还配有总习题,以供读者复习、巩固所学知识.书末附有习题答案与提示,可供教师与学生参考.

本书可作为高等工科类、经济类院校“概率论与数量统计”课

程的教材或教学参考书. 全书共分八章, 第一章至第四章为概率论部分, 需讲授约 32~36 学时, 第五章至第八章为数理统计部分, 其中第五章至第七章需讲授约 22 学时, 第八章需讲授约 12 学时, 各校可根据实际情况与专业需要讲授其中一部分或全部.

本书第一章由茹世才编写, 第二章由欧阳克智编写, 第三、四章由王惠君编写, 第五章至第八章由肖筱南编写. 全书由肖筱南制定编写计划并负责最后统稿定稿.

我们谨将本书奉献给读者, 希望它能成为每位读者学习概率统计的良师益友. 本书的编写与出版得到了北京大学出版社的大力支持与帮助, 在此表示衷心感谢.

限于编者水平和时间仓促, 书中疏漏与不当之处, 恳切读者指正.

编 者
2001 年 10 月

目 录

第一章 随机事件及其概率	(1)
§ 1 随机事件及其运算	(1)
一、随机现象与随机试验	(1)
二、样本空间	(3)
三、随机事件	(4)
四、随机事件间的关系与运算	(5)
习题 1-1	(11)
§ 2 随机事件的概率	(13)
一、概率的统计定义	(13)
二、概率的古典定义	(16)
习题 1-2(1)	(23)
三、概率的几何定义	(25)
四、概率的公理化定义与性质	(28)
习题 1-2(2)	(32)
§ 3 条件概率与全概率公式	(34)
一、条件概率与乘法公式	(34)
二、全概率公式与贝叶斯(Bayes)公式	(38)
习题 1-3	(42)
§ 4 随机事件的独立性	(44)
一、事件的相互独立性	(44)
二、伯努利(Bernoulli)概型及二项概率公式	(49)
习题 1-4	(52)
总习题一	(54)

第二章 随机变量及其分布	(57)
§ 1 离散型随机变量及其分布律	(57)
一、随机变量的定义	(57)
二、离散型随机变量及其分布律	(58)
三、常用的离散型随机变量	(60)
习题 2-1	(65)
§ 2 随机变量的分布函数	(66)
一、分布函数的概念	(67)
二、分布函数的性质	(70)
习题 2-2	(70)
§ 3 连续型随机变量及其概率密度	(71)
一、连续型随机变量的概率密度	(72)
二、连续型随机变量的性质	(74)
三、离散型与连续型随机变量的比较	(78)
习题 2-3	(79)
§ 4 几种常见的连续型随机变量	(81)
一、均匀分布	(81)
二、指数分布	(82)
三、正态分布	(83)
习题 2-4	(89)
§ 5 随机变量函数的分布	(90)
一、离散型情况	(90)
二、连续型情况	(92)
习题 2-5	(95)
§ 6 二维随机变量及其联合分布函数	(96)
一、二维随机变量的概念	(96)
二、联合分布函数的定义及意义	(98)
三、联合分布函数的性质	(99)
习题 2-6	(99)

§ 7	二维离散型随机变量	(100)
	一、联合分布律	(100)
	二、边缘分布律	(102)
	三、条件分布律	(103)
	习题 2-7	(104)
§ 8	二维连续型随机变量	(105)
	一、联合概率密度	(105)
	二、边缘概率密度	(108)
	三、两种重要的二维连续型分布	(109)
	四、条件概率密度	(111)
	习题 2-8	(113)
§ 9	随机变量的相互独立性	(114)
	一、随机变量相互独立的定义	(115)
	二、离散型随机变量独立的充要条件	(115)
	三、连续型随机变量独立的充要条件	(117)
	四、二维正态变量的两个分量独立的充要条件	(117)
	习题 2-9	(118)
§ 10	两个随机变量的函数的分布	(120)
	一、离散型情况	(120)
	二、连续型情况	(121)
	习题 2-10	(126)
	总习题二	(128)

第三章 随机变量的数字特征 (130)

§ 1	数学期望	(130)
	一、离散型随机变量的数学期望	(131)
	二、连续型随机变量的数学期望	(132)
	三、随机变量的函数的数学期望	(134)
	四、数学期望的性质	(137)
	习题 3-1	(139)

§ 2	方差	(140)
	一、方差的定义	(141)
	二、常见分布的方差	(141)
	三、方差的性质	(144)
	习题 3-2	(146)
§ 3	协方差与相关系数	(147)
	一、协方差	(147)
	二、相关系数	(150)
	三、相关系数的意义	(152)
	习题 3-3	(154)
§ 4	矩与协方差矩阵	(154)
	习题 3-4	(156)
	总习题三	(156)
第四章	大数定律与中心极限定理	(158)
§ 1	大数定律	(158)
	习题 4-1	(162)
§ 2	中心极限定理	(163)
	习题 4-2	(166)
	总习题四	(167)
第五章	统计量及其分布	(169)
§ 1	总体与随机样本	(169)
	一、总体与个体	(169)
	二、随机样本与样本值	(170)
	习题 5-1	(172)
§ 2	统计量与抽样分布	(172)
	一、 χ^2 分布	(175)
	二、 t 分布	(177)
	三、 F 分布	(178)
	四、正态总体样本均值与样本方差的抽样分布	(181)

习题 5-2	(184)
§ 3 总体分布的近似描述	(185)
一、样本频数分布与频率分布表	(185)
二、频率直方图	(187)
三、经验分布函数	(190)
习题 5-3	(191)
总习题五	(192)
第六章 参数估计	(194)
§ 1 点估计	(194)
一、矩估计法	(195)
二、顺序统计量法	(197)
三、极大似然估计法	(200)
习题 6-1	(207)
§ 2 估计量的评价标准	(209)
一、无偏性	(210)
二、有效性	(212)
三、一致性	(213)
习题 6-2	(215)
§ 3 区间估计	(217)
习题 6-3	(222)
§ 4 正态总体参数的区间估计	(223)
一、单个总体 $N(\mu, \sigma^2)$ 的情形	(223)
二、两个总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的情形	(227)
习题 6-4	(234)
§ 5 单侧置信区间估计	(236)
习题 6-5	(238)
总习题六	(239)
第七章 假设检验	(243)
§ 1 假设检验的概念与步骤	(243)

一、假设检验的基本概念	(243)
二、假设检验的基本原理与方法	(244)
三、两类错误	(247)
四、假设检验的一般步骤	(248)
习题 7-1	(250)
§ 2 正态总体均值的假设检验	(250)
一、单个总体 $N(\mu, \sigma^2)$ 均值 μ 的检验	(251)
二、两正态总体均值差的检验	(261)
三、基于成对数据的均值差检验	(267)
习题 7-2	(269)
§ 3 正态总体方差的假设检验	(271)
一、单个总体 $N(\mu, \sigma^2)$ 方差 σ^2 的检验	(272)
二、两正态总体方差比的检验	(278)
习题 7-3	(282)
§ 4 总体分布函数的假设检验	(284)
习题 7-4	(290)
总习题七	(291)
第八章 方差分析与回归分析	(295)
§ 1 单因素方差分析	(295)
一、单因素试验	(295)
二、单因素等重复试验的方差分析	(297)
三、不等重复的单因素试验方差分析	(307)
习题 8-1	(309)
§ 2 双因素方差分析	(311)
一、双因素等重复试验的方差分析	(311)
二、双因素无重复试验的方差分析	(318)
习题 8-2	(322)
§ 3 一元线性回归分析	(324)
一、回归分析问题	(324)

二、一元线性回归	(326)
三、可线性化的一元非线性回归	(340)
习题 8-3	(342)
§ 4 多元线性回归分析	(344)
一、回归平面方程的建立	(344)
二、回归平面方程的显著性检验	(346)
习题 8-4	(348)
总习题八	(349)
附表 1 标准正态分布表	(354)
附表 2 泊松分布表	(356)
附表 3 t 分布表	(359)
附表 4 χ^2 分布表	(362)
附表 5 F 分布表	(367)
附表 6 相关系数显著性检验表	(379)
习题答案与提示	(380)

第一章 随机事件及其概率

概率论与数理统计是数学的一个重要分支,它是研究随机现象统计规律性的一门学科,其应用广泛,是科技、管理、经济等工作者的必备数学工具.本章通过随机试验介绍概率论中的基本概念——样本空间、随机事件及其概率,并进一步讨论随机事件的关系及其运算,概率的性质及其计算方法.

§ 1 随机事件及其运算

一、随机现象与随机试验

在自然界及人类社会活动中,所发生的现象是多种多样的,若从其结果能否准确预言的角度去考虑,可分为两大类:一类称为**确定性现象**,另一类称为**随机(或偶然)现象**.

所谓**确定性现象**,是指在一定的条件下必然发生(或必然不发生)的现象.例如,在标准大气压下,把水加热到 100°C ,此时水沸腾是必然发生的现象(而此时水结冰是必然不会发生的现象).只要保持上述条件不变,任何人重复上述试验及进行观察,该现象的结果总是确定的,这类现象的结果是能准确预言的.研究这类现象的数学工具是线性代数、微积分学及微分方程等经典数学理论与方法.

所谓**随机现象**,是指在一定的条件下,可能发生也可能不发生的现象,具有不确定性(或称为偶然性或随机性).例如,掷一枚质地均匀的硬币,其落地后可能是有国徽的一面(称为正面)朝上,也可能是有数字的一面(称为反面)朝上,掷币前不能准确地预言.又如,从含有不合格品的一批某种产品中,任意抽一件检验,其检验

结果可能是合格品,也可能是不合格品,事先无法准确地预言.

通常,人们不论研究何种现象,都离不开对其进行观察(测)或进行实验.为简便起见,我们把对某现象或对某事物的某个特征的观察(测),以及各种各样的科学实验,统称**试验**.为了研究随机现象,同样需要进行试验.这类试验的特征是,在一定的条件下,其试验的可能结果不止一个;一次试验中,可能出现某一结果,也可能出现另一个结果,究竟会出现哪一个试验结果,事先无法准确地预言.对于这类试验,人们实践中发现,就一次试验而言,其试验结果表现出不确定性(偶然性),似乎难以捉摸,但在大量重复试验下,其试验结果却呈现出某种规律性.例如,抛掷硬币试验,一次抛掷,哪一面朝上是随机的(或偶然的),但把同一硬币进行成千上万次抛掷,人们发现,“正面朝上”与“反面朝上”这两个试验结果出现的次数大致各占一半.又例如,从含有不合格品的一批某种产品中,任意抽取一件检验的试验下,“抽到合格品”与“抽到不合格品”两个试验结果都有可能发生,具有随机性(或偶然性),但当重复抽取时,“抽到合格品”的次数与抽取总次数之比却呈现出某种稳定性.随机现象的这种隐蔽着的内在规律性叫做**统计规律性**.本课程的任务就是要研究和揭示这种规律性.

显然,要获得随机现象的统计规律性,必须在相同的条件下,大量重复地做试验.在概率统计中,把这类试验称为**随机试验**,它具有下述三个特性:

- (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,而究竟会出现哪一个结果,在试验前不能准确地预言;
- (3) 试验所有的可能结果在试验前是明确(已知)的,而每次试验必有其中的一个结果出现,并且也仅有一个结果出现.

随机试验简称**试验**,并用字母 E 或 E_1, E_2 等表示.我们就是通过随机试验去研究随机现象的.

例 1 掷一颗骰子,观察出现的点数.这个试验就是一个随机

试验,记为 E_1 . E_1 的所有可能的结果共有 6 个,即“出现 1 点”,“出现 2 点”,…,”出现 6 点”. 试验前无法准确地预言会出现哪一个点数,但试验后必有且仅有其中的一个点数出现.

例 2 将一枚质地均匀的硬币连掷两次,观察出现正、反面的情况. 这里把硬币连掷两次作为一次试验(称为**复合试验**),这也是一个随机试验,记为 E_2 . 若用记号(第一次掷币可能出现的面,第二次掷币可能出现的面)表示试验结果,则 E_2 共有四个可能的结果:

(正面,正面), (正面,反面), (反面,正面), (反面,反面).

例 3 记录电话交换台一小时内接到呼唤的次数. 这是一个随机试验,记为 E_3 . 试验结果(即接到呼唤次数)的可能值为 0, 1, 2, … . 由于难以规定呼唤次数的上限,理论上认为每个正整数都是一个可能的试验结果.

例 4 五件产品中有一件正品(分别记为 Z_1, Z_2, Z_3)和两件次品(分别记为 C_1, C_2),现从中任意取两件,观察取出的产品的正、次品情况. 这也是随机试验,记为 E_4 . E_4 共有 $C_5^2=10$ 个可能的试验结果: $(Z_1, Z_2), (Z_1, Z_3), (Z_2, Z_3), (Z_1, C_1), (Z_1, C_2), (Z_2, C_1), (Z_2, C_2), (Z_3, C_1), (Z_3, C_2), (C_1, C_2)$.

例 5 从一大批某类电子元件中,任意抽取一件,测试其使用寿命(即能正常工作的时间). 这同样是一个随机试验,记为 E_5 . 由于测试的结果是非负实数,并且也难以确定元件寿命的上限是哪一个非负实数,理论上为处理方便,认为任何一个非负实数都是测试的一个可能结果,故该试验的任一可能结果可记为 $t \in [0, +\infty)$.

二、样本空间

研究任何一个随机试验,不仅要搞清该试验所有可能的结果,还要了解它们的含义,而每一个可能的结果的含义是指试验后所观察(测)到的最简单的直接结果,它不包含其余的任何一个可能

的结果. 我们把试验后所观察(测)到的这种最简单每一个直接结果称为该试验的一个**基本事件**. 全体基本事件所构成的集合称为随机试验的**样本空间**. 样本空间通常用字母 Ω (或 U) 表示. 为了区别不同试验的样本空间, 也可以用 Ω_1, Ω_2 等表示. Ω 中的元素即基本事件, 此时称为**样本点**, 常用字母 ω (或 e) 表示, 必要时也可以用 ω_1, ω_2 等表示不同的样本点.

上面例题中, 随机试验的样本空间分别为:

$$E_1: \Omega_1 = \{1, 2, 3, 4, 5, 6\};$$

$$E_2: \Omega_2 = \{(\text{正面}, \text{正面}), (\text{正面}, \text{反面}), (\text{反面}, \text{正面}), (\text{反面}, \text{反面})\};$$

$$E_3: \Omega_3 = \{0, 1, 2, \dots, n, \dots\};$$

$$E_4: \Omega_4 = \{(Z_1, Z_2), (Z_1, Z_3), (Z_2, Z_3), (Z_1, C_1), (Z_1, C_2), (Z_2, C_1), (Z_2, C_2), (Z_3, C_1), (Z_3, C_2), (C_1, C_2)\};$$

$$E_5: \Omega_5 = \{t | t \geq 0\} \text{ 或 } \Omega_5 = [0, +\infty).$$

从上述样本空间中, 可以发现它们中有的是数集, 有的不是数集; 有的数集是有限集, 有的则是无限集.

三、随机事件

当研究随机试验时, 人们通常关心的不仅是某个样本点在试验后是否出现, 而更关心的是满足某些条件的样本点在试验后是否出现. 例如, 在例 5 中, 测试某类电子元件的使用寿命以便确定该批元件的质量. 若假定使用寿命超过 1000 小时为正品, 则人们关心的是试验结果是否大于 1000 小时. 满足这个条件的样本点组成了样本空间的子集. 我们把样本空间的子集称为**随机事件**, 简称**事件**. 事件通常用大写字母 A, B, C 等表示, 必要时也可以用 $A_1, B_1, C_1, A_2, B_2, C_2$ 等表示, 也可以用语言描述加花括号来表示. 例如在例 3 中, $\{\text{呼唤次数不超过 5 次}\}$, 例 4 中 $\{\text{取出的两件产品中恰有一件次品}\}$ 等都是这种表示法. 显然, 基本事件就是仅含一个样本点的随机事件; 一个样本空间, 可以有許多随机事件.

随机试验中,若组成随机事件 A 的某个样本点出现,则称事件 A 发生,否则称事件 A 不发生. 如例 1 中,若用 A 表示{出现奇数点},即{1,3,5},它是 Ω_1 的子集,是一个随机事件,它在一次试验中可能发生,也可能不发生,当且仅当掷出的点数是 1,3,5 中的任何一个时,则称事件 A 发生. 同样,若用 B 表示{出现偶数点},即{2,4,6},它是 Ω_1 的子集,也是一个随机事件.

由于样本空间 Ω 是其本身的一个子集,因而也是一个随机事件,又因为样本空间 Ω 包含所有的样本点,所以每次试验必定有 Ω 中的一个样本点出现,即 Ω 必然发生,因而称 Ω 为**必然事件**. 又因空集 $\emptyset$ 总是样本空间 Ω 的一个子集,所以 $\emptyset$ 也是一个随机事件,由于 $\emptyset$ 不包含任何一个样本点,故每次试验 $\emptyset$ 必定不发生,因而 $\emptyset$ 称为**不可能事件**.

必然事件与不可能事件已无随机性可言,在概率论中,为讨论方便,仍把 Ω 与 $\emptyset$ 当作两个特殊的随机事件.

四、随机事件间的关系与运算

由于事件是样本空间的子集,故事件之间的关系与运算和集合论中集合之间的关系与运算完全类同,但要注意其特有的事件意义.

设 Ω 是给定的一个随机试验的样本空间,事件 $A, B, C, A_k (k=1, 2, \dots)$ 都是 Ω 的子集.

1. 包含关系

若事件 A 发生必导致事件 B 发生,则称事件 B **包含** 事件 A , 或称 A 是 B 的**子事件**,记为 $B \supset A$ 或 $A \subset B$.

在这种情况下,组成事件 A 的样本点都是组成 B 的样本点. 这种包含关系的几何直观如图 1-1 所示.

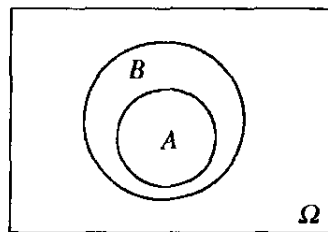


图 1-1 $A \subset B$

例如,在例 1 中,若 A 表示{出现奇数

点},即事件 $\{1,3,5\}$,若 B 表示{出现的点数不超过5},即事件 $\{1,2,3,4,5\}$,显然 $B \supset A$.

2. 相等关系

若 $B \supset A$ 且 $A \supset B$,则称事件 A 与事件 B **相等**,记为 $A=B$.其直观意义是事件 A 与 B 的样本点完全相同.

3. 和事件

{事件 A 与事件 B 至少有一个发生}的事件称为事件 A 与事件 B 的**和事件**,记为 $A \cup B$.事件 $A \cup B$ 是属于 A 或属于 B 的样本点组成的集合,其几何直观由图1-2所示(阴影部分).

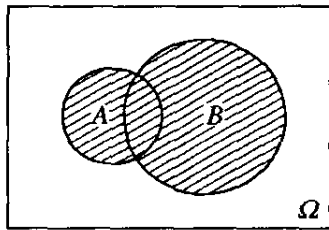


图 1-2 $A \cup B$

例如,在例2中,设 A 表示{第一次掷币出现正面,第二次掷币出现正面},即 $A = \{(\text{正面}, \text{正面})\}$, B 表示{两次掷币恰有一次出现正面},即 $B = \{(\text{正面}, \text{反面}), (\text{反面}, \text{正面})\}$,则 $A \cup B$ 表示{至少有一次出现正面}.

一般地,把{事件 $A_1, A_2, \dots, A_n$ 中至少有一个发生}的事件称为这 n 个事件的和事件,记为 $A_1 \cup A_2 \cup \dots \cup A_n$ 或简记为 $\bigcup_{i=1}^n A_i$.

类似地,把可列个事件即 $\{A_1, A_2, \dots, A_n, \dots\}$ 中至少有一个发生}的事件称为这可列个事件的和事件,记为 $A_1 \cup A_2 \cup \dots \cup A_n \cup \dots$ 或简记为 $\bigcup_{i=1}^{\infty} A_i$.

4. 积事件

{事件 A 与事件 B 同时发生}的事件称为事件 A 与事件 B 的**积事件**,记为 $A \cap B$ 或 AB .事件 $A \cap B$ 是由既属于事件 A 又属于事件 B 的样本点组成的集合,其几何直观由图1-3所示(阴影部分).

例如,在例4中,设 A 表示{取出的两件中最多有一件次品},即 $A = \{(Z_1, Z_2), (Z_1, Z_3), (Z_2, Z_3), (Z_1, C_1), (Z_1, C_2), (Z_2, C_1),$