



新世纪高校统计学专业系列教材

概率论与数理统计

(第 二 版)

GAILÜLUN YU SHULI TONGJI (DI ER BAN)

苏均和 主编



上海财经大学出版社

新世纪高校统计学专业系列教材

概率论与数理统计

(第二版)

苏均和 主编



上海财经大学出版社

图书在版编目(CIP)数据

概率论与数理统计/苏均和主编. —2版. —上海:上海财经大学出版社, 2005. 3

(新世纪高校统计学专业系列教材)

ISBN 7-81049-244-6/F·194

I. 概… II. 苏… III. 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(2005)第 017000 号

☐ 责任编辑 徐 超
☐ 封面设计 优典工作室

GAILÜLUN YU SHULI TONGJI

概率论与数理统计

(第二版)

苏均和 主编

上海财经大学出版社出版发行
(上海市武东路 321 号乙 邮编 200434)

网 址: <http://www.sufep.com>

电子邮箱: webmaster@sufep.com

全国新华书店经销

上海同济大学印刷厂印刷

上海浦东北联装订厂装订

2005 年 3 月第 2 版 2005 年 3 月第 1 次印刷

890mm×1240mm 1/32 13 印张 361 千字

印数 0 001—4 000 定价: 25.00 元

第一版前言

随着我国改革开放形势的不断发展,教育体制以及课程改革也正面临着新的问题。为了使“概率论与数理统计”这门课程适应这种改革形势的需要,我们在1994年12月由中国财政经济出版社出版的《概率论与数理统计》(苏均和主编,财政部统编教材)基础上,从内容体系、案例编排上都做了较大幅度的改变,使这本教材更完善、更成熟。

“概率论与数理统计”是研究随机现象并找出其统计规律的一门学科,是广泛应用于社会、经济、科学等各个领域的定量和定性分析的科学体系。我们要求,通过学习该课程使读者掌握概率、统计的基本概念,熟悉数据处理、数据分析、数据推断的各种基本方法,并能用所掌握的方法具体解决社会经济所遇到的各种问题。

本书共分九章。第一部分(第一至第四章)概率论,主要讲述概率基础,包括:随机变量及其分布、随机变量的数字特征、大数定律及中心极限定理等内容。第二部分(第五至第九章)数理统计,阐述数理统计基础、参数估计、假设检验、非参数方法和回归分析等内容。

这本教材有以下特点:①具有财经特色,所选的案例尽可能与社会经济相结合;②考虑计算机对统计的重要作用,我们在有些统计方法介绍后附有计算机程序;③适用面广,这本教材既可作为统计专业本科的教材,又可作为各类读者的统计入门教材;④理论和方法相结合,这本教材既注重基础理论阐述而且更重视各类方法的具体使用。

本书由苏均和主编,王学民、朱鸣雄、黄国群参加了编写。徐国祥

教授对这本教材的编写、出版,自始至终给予了热情关心和帮助,在此表示感谢。

我们力求使这本教材完善和成熟,但限于编者的水平,难免会有错误,敬请读者斧正。

编 者

1998年5月12日

第二版前言

本教材自1994年以财政部部属院校统编教材的形式问世以来,已经历了十余年时间。承读者厚爱,它经过两家出版社五次修改重印,读者也增至数万人。它迄今仍能受到欢迎,一方面是由于教材内容适应当今形势的发展,另一方面也说明教材本身满足了财经类院校发展的需要,反映了教材具有较强的生命力。

然而该教材在各院校使用过程中和读者在学习过程中都反映各种不足之处。其一是原教材在内容、符号等方面都有一定程度的错误,使学生在在学习上发生疑惑和困难;其二是内容、习题编排上不能适应学生更高的要求。面对各院校反映的情况和读者来信所提出的问题,我们认为很有必要对此教材进行一次较大幅度的修正和调整。为此,组织人员做了以下几项工作:

一、收集整理各院校及读者来信,分门别类地分析他们所提出的问题;

二、分层次(即让本科生、研究生及相关教师)通读教材数遍以上,以指出它的错误和不足;

三、根据读者提出的各类意见和想法,提出此书的修改大纲和内容增删计划,最终得到完整的框架和新教材的编写思路。

众所周知,“概率论与数理统计”课程是一门财经类院校的重要基础课程。它在深化改革及服务于现今形势发展所起的作用是毫无疑问的。因此我们以极其认真甚至是神圣的态度重编了这本教材,希

望能得到读者进一步认可和欢迎。尽管如此,本书还可能会发生新的问题,敬请批评指正。

编 者

2004 年 11 月 20 日

目 录

第一章 随机事件和概率	1
第一节 随机事件及其运算	1
第二节 随机事件的概率及其性质	8
第三节 条件概率、乘法公式、全概率公式与贝叶斯公式	21
第四节 相互独立的随机事件与独立试验概型	28
本章小结	32
习题一	33
第二章 随机变量及其分布	38
第一节 随机变量及其分布函数	38
第二节 离散型随机变量	42
第三节 连续型随机变量	53
第四节 多维随机变量及其独立性	66
第五节 随机变量函数的分布	79
本章小结	90
习题二	91
第三章 随机变量的数字特征和特征函数	99
第一节 随机变量的数学期望	99
第二节 随机变量的方差、协方差与相关系数	111

第三节 随机变量的特征函数	128
本章小结	133
习题三	134
 第四章 大数定律和中心极限定理	 139
第一节 大数定律	139
第二节 中心极限定理	145
本章小结	157
习题四	159
 第五章 数理统计的基本概念	 161
第一节 总体与样本	161
第二节 统计量与样本矩	163
第三节 抽样分布	166
本章小结	178
习题五	178
 第六章 参数估计	 181
第一节 参数的点估计	181
第二节 点估计量的优良标准	190
第三节 参数的区间估计	202
本章小结	214
习题六	215
 第七章 假设检验	 220
第一节 假设检验的基本概念	220
第二节 单个正态总体下参数的检验	224
第三节 两个正态总体下参数的比较检验	232
第四节 非正态总体下的大样本参数检验	237

本章小结	242
习题七	243
第八章 非参数统计推断	245
第一节 分布的 χ^2 -检验	245
第二节 柯尔莫哥洛夫和斯米尔诺夫检验	255
第三节 偏度、峰度检验方法与正态概率纸方法	262
第四节 符号检验与中位数检验	267
第五节 游程检验与秩和检验	272
本章小结	278
习题八	278
第九章 回归分析	281
第一节 概述	281
第二节 一元回归分析和相关分析	282
第三节 非线性回归	293
第四节 多元线性回归	299
第五节 逐步回归	314
本章小结	316
习题九	317
补充习题	322
附录一 习题参考答案及提示	348
附录二 补充习题参考答案及提示	364
附录三 关于上、下、双侧分位数	377

附录四 统计用表	380
1. 随机数字表	380
2. 正态分布双侧临界值表	381
3. 二项分布表	382
4. 泊松分布表	384
5. 正态分布函数 $N(0,1)$ 的数值表	386
6. t -分布双侧临界值表	388
7. χ^2 -分布的上侧临界值 χ^2_α 表	389
8. F -分布上侧临界值表	390
9. 柯尔莫哥洛夫检验的临界值 (D_m^*) 表	398
10. 秩和检验表	399
11. 相关系数检验表	400
12. 符号检验表	401
13. 游程数检验表	402

第一章 随机事件和概率

第一节 随机事件及其运算

一、必然现象和随机现象

在自然界、生产实践、科学试验和日常生活中,人们观察到的现象大体可以归结为两种类型。一类是在一定条件下必然出现(或者不出现)某种结果的现象,这类现象的一个共同特点就是可以事前预言,即在准确地重复某些条件下,它的结果总是可以肯定的,或是根据它过去的状态,在相同的条件下完全可以预言将来的发展,我们把这类现象称之为确定性现象或者必然现象。例如,向上抛掷的重物必然自由下落;直角三角形的三边必然满足勾股定理;在一批合格的产品中任意取一件,必定不是废品;等等。几何、微积分、线性代数都是研究确定性现象的数学工具。另一类现象是在相同的条件下可能得到多种不同的结果,出现具体哪一种结果是不可预测的,即在相同的条件下,未来的发展事前不能肯定。例如,将一枚硬币抛向空中,无论如何控制硬币的抛掷条件,我们在每次抛掷前总是无法肯定硬币自由下落后,我们可以看见硬币的哪一面;又如某射手向同一目标多次射击,各次弹着点不尽相同,并且无论事前怎样控制射击条件,在一次射击之前无法肯定弹着点的确切位置。这类现象的共同特点就是:可以在相同条件下重复进行试验或者观察,而每次试验或者观察的可能结果不止一个,且事前不能预知确切结果,即试验结果呈现出不确定性。人们经过长期实践并深入研究之后,发现这类现象虽然在个别试验或观察中,其结果呈现出不确定性,但是在大量观察或多次重复试验后,其结果往往呈现出某种客

观规律,并且这种客观规律性是可以认识的。例如,多次重复抛掷同一枚均匀硬币,正面朝上的次数大致占抛掷总次数的一半;某射手向同一目标射击的弹着点按照一定的规律分布;等等。这种在大量重复试验或者观察中所呈现出事物的固有规律称为“统计规律性”。我们把在个别试验或者观察中呈现的不确定性,在大量重复试验或者观察中又具有统计规律的现象称为随机现象(或者偶然现象)。

概率论与数理统计就是研究随机现象统计规律性的一门数学学科,但是各有侧重。概率论侧重于理论上的研究,介绍随机现象反映的基本概念,建立相应的定理和公式,找出计算统计规律的方法;而数理统计是以概率论为理论依据,研究如何设计试验并对试验结果进行整理和统计分析。

概率论的研究开始于意大利文艺复兴时期,源于赌博。数理统计学是伴随着概率论而发展的,最早见于国家的人口统计中。当今,概率论的理论和方法几乎遍及所有的科学技术领域以及工农业生产和国民经济各个部门之中。在经济科学中,它是经济管理、质量控制、保险理论、系统论、经济预测和决策、计量经济学等的理论基础,是各类经济、管理专门人才不可缺少的数学工具。

二、随机试验与随机事件、样本空间

为了叙述方便,我们把对某种自然现象作一次观察或者进行一次科学试验,统称为一个试验。如果这个试验具有下列特性:

- (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,并且能在试验之前明确知道所有可能的结果。
- (3)每次试验之前不能肯定这次试验会出现哪一个结果,但是可以肯定每次试验总会出现并且仅出现这些可能结果中的一个。

我们将具有上述三个特性的这种试验称为随机试验,简称为试验。在随机试验中,我们所关心的是某种结果是否出现,把试验的每一个可能结果称之为随机事件,简称事件。

通过试验观察随机现象时,假定试验 E 中一切可能的结果是已经知道的,并且每次试验都必定是其中一个结果出现,我们把这些结果组成的集合称为样本空间或者基本事件空间,记做 Ω 。 Ω 的元素(即试验中的一个可能结果)称为样本点或基本事件,记做 ω 。

在研究随机现象时,不仅考虑基本事件(即不可能再分的事件),而且还要考虑一些由基本事件组成的复合事件。一般把事件定义为样本空间 Ω 的子集,用大写字母 A, B 等来表示。在试验中,当且仅当事件 A 中所包含的某一样本点出现时,称事件 A 发生。

下面举一些样本空间和事件的例子。

例 1 做掷一枚均匀硬币的试验 E ,用“正”表示“出现正面”,用“反”表示“出现反面”,试验的可能结果只有两个: $w_1 = \text{“正”}$, $w_2 = \text{“反”}$,试验 E 的样本空间 $\Omega = \{w_1, w_2\}$ 由两个简单事件组成。

例 2 设试验 E 为同时掷两枚可以区别的硬币,试验的可能结果共有四个: $w_1 = \{\text{正, 正}\}$, $w_2 = \{\text{正, 反}\}$, $w_3 = \{\text{反, 正}\}$, $w_4 = \{\text{反, 反}\}$,试验 E 的样本空间 $\Omega = \{w_1, w_2, w_3, w_4\}$ 。

在这个试验中,还可以考虑它的某些复合事件,例如,

A:“两个硬币出现同一面”的事件,则事件 A 由 w_1, w_4 组成。即 $A = \{w_1, w_4\}$ 。这时,当且仅当 w_1 或者 w_4 之一出现时,事件 A 就发生了。

B:“恰有一个硬币出现正面”的事件,则事件 $B = \{w_2, w_3\}$ 。

C:“至少有一个硬币出现正面”的事件,则事件 $C = \{w_1, w_2, w_3\}$,这时,当且仅当 w_1, w_2 或者 w_3 之一出现时,事件 C 就发生了。

例 3 某电话交换台,在单位时间内接到呼唤的次数是 $0, 1, 2, \dots$,用 ω_i 表示接到 i 次呼唤的事件,并简记 ω_i 为 i ,则样本空间 $\Omega = \{0, 1, 2, \dots\}$ 。

在这一试验中,可以考虑 $A = \{i | i > 200, i \text{ 为正整数}\}$, $B = \{i | i > 500, i \text{ 为正整数}\}$ 等等这样一些复合事件。

例 4 在一批灯泡中,任意取一只测试其寿命 $t(\text{h})$,则样本空间为 $\Omega = \{t | t > 0\}$ 。

在这一试验中,可以考虑 $A = \{t | t > 500\}$, $B = \{t | t < 2\,000\}$ 等等这样一些复合事件。

由上面例子可知,对于一个具体的试验来说,样本空间是根据试验的具体内容来决定的。样本空间可以相当简单,也可以相当复杂。但是在概率论的研究讨论中,一般认为样本空间是给定的(或者是事前已经明确了试验的所有可能的结果)。这种必要的抽象,其目的是为了能够更好地把握随机现象的本质,而使所得的理论性的结论得到更加广泛的应用。当然,对一个实际问题来讲,样本空间如何描述,应进行具体的分析和研究。此外,一个抽象的样本空间经常可以被赋予各种不同的解释(即可以概括各种内容不同的问题)。例如,仅仅包含两个样本点的样本空间,它既可以作为掷硬币出现“正面”或者出现“反面”的模型,也能用于检验产品的“合格”或者“不合格”;用于公用事业中排队现象的“有人排队”或“无人排队”;以及出生婴儿性别的“男”或“女”;等等。这说明尽管问题的实际内容不同,但是有时候却能够归结为相同的概率模型。因此,我们常常以抛掷硬币、摸球等这样的一些既典型又形象且易于理解的例子来阐明一些问题,以便使问题阐述得更加明确,使问题的本质更为突出。

为了研究的方便,把样本空间 Ω 也作为一个事件。因为在每次试验时,必然出现 Ω 中的某个样本点,即 Ω 必然发生,所以也称 Ω 为必然事件,类似地空集 Φ 也作为一个事件,它在每次试验中都不会发生,故称为不可能事件。

三、事件的关系与运算

1. 事件的包含与相等

如果任一属于 A 的样本点一定属于 B ,则称事件 B 包含事件 A ,记为 $B \supset A$,或者 $A \subset B$ 。显然, $B \supset A$ 的含义是,事件 A 发生,必然导致事件 B 发生。

例如,以 A 表示“灯泡使用寿命大于 500 小时”,以 B 表示“灯泡使用寿命大于 200 小时”,则 $B \supset A$ 。

如果 $B \supset A$ 与 $A \supset B$ 同时成立,则 A 与 B 相等,记作 $A=B$ 。

2. 事件的并

由属于 A 或者属于 B 的所有样本点组成的集合,称为 A 与 B 的

并(或者和),记为 $A \cup B$ 或者 $A+B$ 。显然,事件 $A \cup B$ 表示事件 A 与事件 B 至少有一个发生。

3. 事件的交

由属于 A 同时又属于 B 的所有样本点组成的集合,称为 A 与 B 的交(或者积),记为 $A \cap B$ 或者 AB 。显然,事件 $A \cap B$ 表示事件 A 与事件 B 同时发生。

4. 事件的差

由属于 A 但是不属于 B 的所有样本点组成的集合,称为 A 与 B 的差,记为 $A-B$ 。事件 $A-B$ 表示事件 A 发生而事件 B 不发生。显然, $B-A$ 表示事件 B 发生而事件 A 不发生。

5. 对立事件

样本空间 Ω 与 A 的差 $\Omega-A$ 称为事件 A 的对立事件(或者逆事件),记为 $\bar{A}$,事件 $\bar{A}$ 表示事件 A 不发生。

例如,“灯泡使用寿命大于 500 小时”的事件是“灯泡使用寿命不大于 500 小时(小于等于 500 小时)”事件的对立事件。

显然, A 与 $\bar{A}$ 互为对立事件,即有 $\bar{\bar{A}}=A$ 。

对立事件 A 和 $\bar{A}$ 间的关系可以表示为: $A \cup \bar{A} = \Omega$, $A \cap \bar{A} = \Phi$ 。

6. 互不相容事件

如果 $A \cap B = \Phi$,则称事件 A, B 互不相容。

基本事件是互不兼容的。

事件的并与事件的交的概念可以推广到有限多个事件或者可列多个事件的场合。

注 $\bigcup_{k=1}^n A_k$ 表示 $A_1 \cup A_2 \cup \cdots \cup A_n$, $\bigcap_{i=1}^n A_i$ 表示 $A_1 \cap A_2 \cap \cdots \cap A_n$,下同。

称 $\bigcup_{k=1}^n A_k$ 为事件 $A_1, A_2, \dots, A_n$ 的并,通常表述为“ $A_1, A_2, \dots, A_n$ 至少有一个发生”。

称 $\bigcup_{k=1}^{\infty} A_k$ 为事件 $A_1, A_2, \dots, A_n, \dots$ 的并,通常表述为“ $A_1, A_2, \dots, A_n, \dots$ 至少有一个发生”。

称 $\bigcap_{k=1}^n A_k$ 为事件 $A_1, A_2, \dots, A_n$ 的交, 通常表述为“ $A_1, A_2, \dots, A_n$ 同时发生”。

称 $\bigcap_{k=1}^{\infty} A_k$ 为事件 $A_1, A_2, \dots, A_n \dots$ 的交, 通常表述为“ $A_1, A_2, \dots, A_n \dots$ 同时发生”。

由上面的定义可以看出事件的关系和运算与集合的相应关系和运算一致, 因此也常常用图形来直观地表示, 见图 1-1。

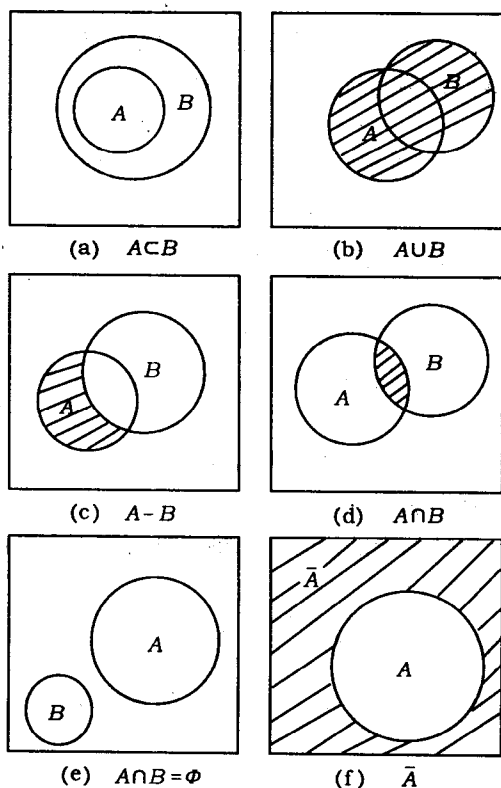


图 1-1 事件的关系与运算