普通高等教育“十三五”规划教材

概率论与数理统计教程

刘金山　主编

科学出版社

普通高等教育"十三五"规划教材

概率论与数理统计教程

刘金山　主编

科学出版社

北京

内 容 简 介

本书共 12 章. 第 1~5 章为概率论部分, 第 6~10 章为数理统计部分, 第 11 章是贝叶斯估计, 第 12 章是 R 软件简介.

本书参照教育部教学指导委员会制定的非数学类概率论与数理统计课程教学基本要求, 结合编者多年来的教学体会, 在对已有教材进行改进的基础上编写而成. 本书特点是论述严谨、通俗易懂、注重应用, 力求深入浅出, 便于学生学习掌握概率论与数理统计的基本内容和方法, 并了解和掌握一些现代统计方法及软件应用.

本书适合普通高等学校非数学、非统计学类各专业概率论与数理统计课程教材或学习参考书, 特别是比较适合工科、理科、经济、管理和农林类各专业, 也可作为各类科技和管理人员的参考书.

图书在版编目(CIP)数据

概率论与数理统计教程/刘金山主编. —北京: 科学出版社, 2016
普通高等教育"十三五"规划教材
ISBN 978-7-03-048628-8

Ⅰ. ①概… Ⅱ. ①刘… 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 126685 号

责任编辑: 姚莉丽 张中兴 / 责任校对: 张凤琴
责任印制: 白 洋 / 封面设计: 迷底书装

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

安泰印刷厂 印刷
科学出版社发行 各地新华书店经销
*
2016 年 6 月第 一 版 开本: 720 × 1000 1/16
2016 年 6 月第一次印刷 印张: 21.5
字数: 431 000

定价: 43.00 元
(如有印装质量问题, 我社负责调换)

《概率论与数理统计教程》编委会

主　编　刘金山

副主编　李泽华　肖　莉

参　编　夏　强　利小玲

杨志程　李　朗

前　言

本书参照教育部非数学类专业数学基础课程教学指导委员会制定的非数学类专业概率论与数理统计课程教学基本要求, 结合多年来教学实践中的经验和体会, 在对已有教材进行认真改进的基础上编写而成, 其目的是为高等学校工科、理科、经济、管理和农林类各专业学生提供一本比较适合的教材或学习参考书.

概率论与数理统计是定量地研究随机现象统计规律的现代数学分支之一, 有着非常广泛的应用背景, 在工业、农业、商业、军事、科学研究、工程技术、经济管理等几乎所有领域都有重要应用. 随着现代科学技术的迅猛发展, 特别是计算机和信息技术的发展, 近年来概率统计方法在经济、金融、保险、生物、农林、医学和管理等许多领域中得到了广泛应用和深入发展. 正是这种广泛的应用性, 使得概率论与数理统计课程成为高等学校大部分专业开设的一门重要的必修或选修课程. 通过本课程的学习可以使学生学习掌握处理随机性观察数据的基本理论和方法, 为各专业知识的深入学习或应用打下良好的基础.

关于概率论与数理统计的教材或教科书已非常多, 这类教材主要以经典概率统计的理论为基础, 讲述其理论、方法与例题分析, 目的是帮助读者理解和掌握基本的概率统计概念和方法. 对于统计计算, 由于分析的基本是少量的数据, 一般经过简单手工计算就可以解决问题. 随着信息技术的快速发展, 大数据时代的来临, 目前在科学技术和日常生活中需要解决的概率统计问题涉及的数据量往往非常大, 人们需要采用概率统计方法分析和解决真实的数据时, 仅仅靠这些传统和经典的内容、方法和手段就显得不够了, 这就为概率统计的学科发展和教学内容的改革提出急迫的挑战. 为此, 本书力求在以下五个方面做一些尝试:

(1) 以概率论为基础, 以数理统计为主线, 立足于概率统计的基本理论和方法;

(2) 尽量与现代科学技术, 特别是信息技术发展相适应, 强调应用性、实效性;

(3) 兼顾传统结构体系和现代方法技术, 适当加强现代贝叶斯方法和 R 软件技术;

(4) 有一定的可塑性, 能广泛适用于普通大学理科、工科以及经济类、管理类、农林类等专业, 各类专业的学生可根据其特点和需要选择教学内容和习题;

(5) 深入浅出, 易教易学, 突出重点, 强调案例式教学方法.

当然, 上述想法只是编者编写本书的希望或初衷, 本书实际上还远没有达到这样的目标.

本书共分 12 章, 内容包括: 随机事件及其概率, 一维随机变量及其分布, 多维

随机变量及其分布, 随机变量的数字特征, 极限定理, 抽样分布理论, 参数估计, 假设检验, 方差分析, 回归分析, 贝叶斯估计, 以及 R 软件简介. 各章包含大量例题和习题, 有些内容还提供了用 R 软件进行统计分析的程序和示例, 书末配有习题参考答案和附表.

本书第 1~5 章初稿由刘金山和肖莉执笔, 第 6~8 章初稿由李泽华执笔, 第 9 章初稿由利小玲执笔, 第 10 章初稿由杨志程执笔, 第 11 章初稿由夏强执笔, 第 12 章初稿由李朗执笔. 刘金山负责全书的统稿和定稿.

由于编者水平有限, 书中难免有缺点和错误, 敬请读者批评指正.

编　者

2015 年 12 月

目　录

第 1 章　随机事件及其概率

在自然界和人类社会活动中, 人们所观察的现象大致上可分为两类. 一类是事先可以预知结果的现象, 即在一定条件下, 某一确定的现象必然会发生, 或根据它过去的状态, 完全可以预知它将来的发展状态. 我们称这一类现象为**确定性现象**或**必然现象**. 例如, 在一个标准大气压下, 水加热到 100 ℃时必然沸腾. 另一类是事先不能预知结果的现象, 即在相同条件下重复进行试验或观测时, 每次出现的结果未必相同, 或者即使知道它过去的状态, 也不能完全确定它将来的发展状态. 我们称这一类现象为**随机现象**或**偶然现象**. 例如, 多次抛掷一枚骰子, 朝上一面出现的点数可能是 $1,2,\cdots,6$ 中的任何一个, 但在每次抛掷前不能预知出现的点数到底是几. 这类现象的共同特点是：在相同条件下重复进行试验或观测, 其结果不止一个, 在每次试验之前不能预知该次试验的确切结果.

对于随机现象, 人们通过大量的实践发现, 在相同的条件下, 虽然试验结果在一次试验或观察中到底出现哪个是不确定的, 但在大量重复试验中却能呈现出某种规律性, 这种规律性称为统计规律性. 例如, 多次抛掷一枚均匀的硬币时, 带国徽的一面朝上的次数约占总抛掷次数的一半.

概率论与数理统计就是以数量化方法研究随机现象统计规律的学科. 概率论是研究随机现象的模型, 即概率分布; 数理统计是研究随机现象的数据分析与处理方法. 概率论与数理统计不仅研究能大量重复的随机现象, 而且也研究不能重复的随机现象, 例如, 某些经济现象 (如经济增长速度、金融产品收益率、股票价格等).

1.1　基 本 概 念

1.1.1　随机试验与事件

在概率论与数理统计中, “试验” 是一个广泛的术语. 我们把在一定条件下对某种现象的一次观察、测量或进行一次科学实验, 统称为一个试验. 一般称满足下面两个条件的试验为**随机试验**:

(1) 在相同条件下可以重复进行;

(2) 每次试验结果事先不可预知, 但所有可能的试验结果事先知道.

一般用字母 E 表示随机试验. 下面是一些随机试验的例子：

E_1：抛掷一颗骰子, 观察出现的点数;

E_2：将一枚硬币连续抛掷两次, 观察其正反面出现的情况;

E_3：将一枚硬币连续抛掷两次, 观察其正面出现的次数;

E_4：观察一天内进入某个超市的顾客人数;

E_5：观察某型号电视机的使用寿命 t;

E_6：记录某地区一昼夜的最低气温 x 和最高气温 y.

对一个随机试验, 把所有可能的试验结果组成的集合称为该试验的**样本空间**, 记为 Ω. 样本空间中的每个元素称为**样本点**. 在上述 6 个试验中, 若以 Ω_i 表示试验 E_i 的样本空间, 则

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

$\Omega_2=\{HH,HT,TH,TT\}$, 其中 H 表示正面, T 表示反面;

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

$\Omega_4=\{0,1,2,\cdots\}$;

$\Omega_5=\{t|0\leqslant t<\infty\}$;

$\Omega_6=\{(x,y)|T_0\leqslant x\leqslant y\leqslant T_1\}$, 其中 T_0 和 T_1 分别表示这一地区的最低气温和最高气温.

对于样本空间应注意下面三点:

(1) 样本空间是一个集合, 它由样本点组成, 可以用列举法或描述法来表示;

(2) 在样本空间中, 样本点可以是一维的, 也可以是多维的, 样本点个数可以是有限的, 也可以是无限的;

(3) 对于一个随机试验而言, 试验的目的不同, 样本空间往往也不同. 例如, E_2 和 E_3 虽然都是将一枚硬币抛掷两次, 但由于试验目的不同, 因此样本空间不同, E_2 的样本空间为 $\Omega_2=\{HH,HT,TH,TT\}$, E_3 的样本空间为 $\Omega_3=\{0,1,2\}$.

我们把样本空间的任一个子集称为一个**随机事件**, 简称为事件, 常用大写字母 $A,B,C,\cdots$ 表示. 因此, 随机事件就是随机试验的某些结果 (样本点) 组成的集合. 特别是, 由一个样本点组成的单点集合称为**基本事件**. 在一个试验中, 事件 A 发生当且仅当 A 中某个样本点出现, 这就是事件 A 发生的含义.

例 1.1.1　在抛掷一颗骰子的试验中, 若用 A 表示 “出现偶数点”, B 表示 “出现奇数点”, C 表示 “出现 3 点或 3 点以上”. 假设试验的目的是观察出现的点数, 试用样本点表示事件 A,B,C.

解　事件 A,B,C 分别表示为 $A=\{2,4,6\}$, $B=\{1,3,5\}$, $C=\{3,4,5,6\}$.

例 1.1.2　从一批电脑中任取一台, 观察其无故障运行的时间 T(单位：小时). A 为事件 “恰好运行 240 小时”, B 为事件 “运行 240.2 小时以上”, C 为事件 “运行时间大于 270.5 小时, 小于等于 480.7 小时”. 试写出样本空间及事件 A,B,C.

解　样本空间为 $\Omega=\{T|T\geqslant 0\}$, 事件 A,B,C 分别为 $A=\{T=240\}$, $B=\{T|T>240.2\}$, $C=\{T|270.5<T\leqslant 480.7\}$.

样本空间 Ω 是其自身的一个子集, 因此它也是一个事件, 由于它包含所有样本点, 所以在每次试验中它必然发生, 因此 Ω 表示必然事件. 空集 $\varnothing$ 不包含任何样本点, 每次试验它都不会发生, 故 $\varnothing$ 表示不可能事件. 虽然它们不是真正的随机事件, 但为了研究问题的方便, 我们把它们视为特殊的随机事件.

1.1.2 事件的关系与运算

因为事件是集合, 即样本空间的子集, 所以事件之间的关系和运算可以按照集合之间的关系和运算来处理. 根据 "事件发生" 的含义, 不难给出事件的关系与运算的定义和规则.

设 Ω 是样本空间, A, B, C 及 $A_1, A_2, \cdots$ 都是事件, 即 Ω 的子集, 它们有以下关系.

(1) **包含关系**. 若 A 的发生必然导致 B 的发生, 则称 B 包含 A 或 A 是 B 的子事件, 记为 $B \supset A$ 或者 $A \subset B$, 即 A 的元素全属于 B(图 1.1.1).

(2) **相等关系**. 若 $A \subset B$ 且 $B \subset A$, 则称 A 与 B 相等, 记为 $A = B$.

(3) **事件的和**. 对两个事件 A 和 B, 定义事件

$$C = \{A\text{发生或}B\text{发生}\},$$

称其为 A 与 B 的和事件, 记为 $C = A \cup B$. 事件 $A \cup B$ 发生, 即 A 发生或 B 发生, 意味着 A 与 B 至少有一个发生 (图 1.1.2).

和事件可以推广到多个事件的情形. 设有 n 个事件 $A_1, A_2, \cdots, A_n$, 定义它们的和事件为

$$C = \bigcup_{k=1}^{n} A_k = \{A_1, A_2, \cdots, A_n\text{中至少有一个发生}\}.$$

对无穷多个事件 $A_1, A_2, \cdots, A_n, \cdots$, 可以类似地定义它们的和事件为

$$C = \bigcup_{k=1}^{\infty} A_k = \{A_1, A_2, \cdots, A_n, \cdots\text{中至少有一个发生}\}.$$

(4) **事件的积**. 对两个事件 A 和 B, 定义事件

$$C = \{A\text{发生且}B\text{发生}\},$$

称其为 A 与 B 的积事件, 记为 $C = A \cap B$(或 $C = AB$). 事件 AB 发生意味着 A 发生且 B 发生, 即 A 与 B 同时发生 (图 1.1.3).

类似地, 可以定义多个事件 $A_1, A_2, \cdots, A_n$ $\cdots$ 的积事件. 根据事件的个数为有限和无限情况分别有下列积事件:

$$C = \bigcap_{k=1}^{n} A_k = \{A_1, A_2, \cdots, A_n\text{同时发生}\},$$

$$C=\bigcap_{k=1}^{\infty}A_k=\{A_1,A_2,\cdots,A_n,\cdots\text{同时发生}\}.$$

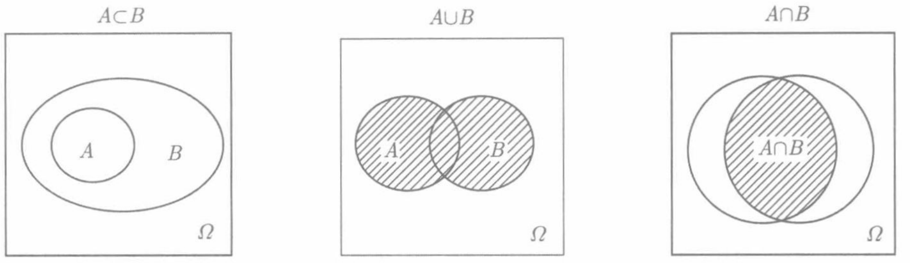

图 1.1.1　　图 1.1.2　　图 1.1.3

(5) **事件的差**. 对两个事件 A 和 B, 定义事件

$$C=\{A\text{发生且}B\text{不发生}\},$$

称其为 A 与 B 的差事件, 记为 $C=A-B$ (或 $C=A\backslash B$), 即 A 发生但 B 不发生的事件 (图 1.1.4). 容易知道 $A-B=A-AB$.

(6) **互斥事件**. 若两个事件 A 与 B 不能同时发生, 即 $AB=\varnothing$, 则称 A 与 B 是互斥事件, 或称它们互不相容 (图 1.1.5). 若事件 $A_1,A_2,\cdots,A_n$ 中任意两个都互斥, 则称事件组 $A_1,A_2,\cdots,A_n$ 两两互斥.

当事件 A 与 B 互斥时, 可记它们的和事件 $A\cup B$ 为 $A+B$; 对于两两互斥的多个事件的和事件有类似的记法.

(7) **对立事件**. “A 不发生” 的事件称为 A 的对立事件, 记为 $\overline{A}$, 即 $\overline{A}=\Omega-A$(图 1.1.6), 并称 A 与 $\overline{A}$ 为互逆事件, 它们是互为对立的事件, 满足 $A\cup\overline{A}=\Omega$, $A\overline{A}=\varnothing$, $\overline{\overline{A}}=A$.

例如, 在抛掷硬币的试验中, 设 A 为 “出现正面”, B 为 “出现反面”, 则 A 与 B 互斥且 A 与 B 互为对立; 在掷骰子的试验中, 设 A 为 “出现 1 点”, B 为 “出现 3 点以上”, 则 A 与 B 互斥, 但 A 与 B 不是对立事件.

利用事件对立关系容易知道, 对于任意事件 A 和 B, $A-B=A\overline{B}$.

设 A,B,C 为事件, 根据集合的运算规则, 有以下事件运算规则.

(1) **交换律**　$A\cup B=B\cup A$; $AB=BA$.

(2) **结合律**　$(A\cup B)\cup C=A\cup(B\cup C)$; $(AB)C=A(BC)$.

(3) **分配律**　$A(B\cup C)=(AB)\cup(AC)$; $A\cup(BC)=(A\cup B)(A\cup C)$.

(4) **对偶律**　$\overline{A\cup B}=\overline{A}\ \overline{B}$; $\overline{AB}=\overline{A}\cup\overline{B}$.

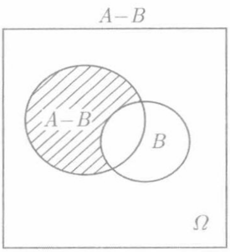

图 1.1.4

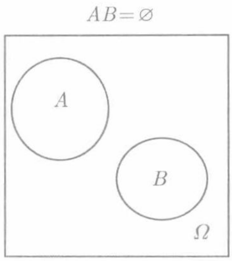

图 1.1.5

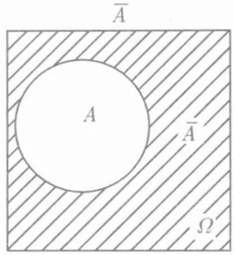

图 1.1.6

对于多个事件情况, 上述运算规则仍然成立. 例如:

$$
\begin{aligned}
&A(A_1 \cup A_2 \cup \cdots \cup A_n) = (AA_1) \cup (AA_2) \cup \cdots \cup (AA_n);\\
&A \cup (A_1A_2 \cdots A_n) = (A \cup A_1)(A \cup A_2) \cdots (A \cup A_n);\\
&\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \overline{A}_1\overline{A}_2 \cdots \overline{A}_n;\\
&\overline{A_1A_2 \cdots A_n} = \overline{A}_1 \cup \overline{A}_2 \cup \cdots \cup \overline{A}_n.
\end{aligned}
$$

上述运算规则也可以推广到无穷多个事件的情况.

例 1.1.3 向指定目标连续射击三次, 观察击中目标的情况. 分别用 A_1, A_2, A_3 表示事件 “第一、二、三次射击时击中目标”, 试用 A_1, A_2, A_3 表示以下各事件:

(1) 只第一次击中;

(2) 只击中一次;

(3) 三次都未击中;

(4) 至少击中一次.

解 (1) 事件 “只第一次击中”, 意味着第二次和第三次都不中. 所以, 该事件可表示为 $A_1\overline{A}_2\overline{A}_3$.

(2) 事件 “只击中一次”, 并不指定哪一次击中. 意味着三个事件 “只第一次击中”“只第二次击中” 和 “只第三次击中” 至少有一个发生, 即它们的和事件发生. 由于上述三个事件两两互斥, 所以, 该事件可表示为 $A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3$.

(3) 事件 “三次都未击中”, 就是事件 “第一、二、三次都未击中”, 该事件可表示为 $\overline{A}_1\overline{A}_2\overline{A}_3$ 或 $\overline{A_1 \cup A_2 \cup A_3}$.

(4) 事件 “至少击中一次”, 就是事件 “第一、二、三次射击中至少有一次击中”, 所以, 该事件可表示为 $A_1 \cup A_2 \cup A_3$.

1.1.3 事件域

我们把事件 A 定义为样本空间 Ω 的子集, 但一般来说并不把 Ω 的一切子集作为事件类, 因为这将给定义概率带来困难. 譬如, 当样本空间是实数轴上的一个区间时, 可以人为地构造出无法测量其长度的子集, 如果将这些也看成是事件, 那