

概率论与数理统计

■ 梅国平 袁捷敏 编 著 ■
毛小兵 胡长和

中国商业出版社

面向 21 世纪高校新编教材

C21.43
M44

概率论与数理统计

梅国平 袁捷敏 编著
毛小兵 胡长和

中国商业出版社

图书在版编目(CIP)数据

概率论与数理统计:概率论与数理统计/梅国平等编著,北京:
中国商业出版社,2001.3

ISBN 7-5044-4047-7

I. 概… II. 梅… III. ①概率论-应用-经济管理②数理
统计-应用-经济管理 IV. F224.7

中国版本图书馆 CIP 数据核字(2001)第 11982 号

责任编辑:马一波

封面设计:程其铝

中国商业出版社出版发行
(100053 北京广安门内报国寺1号)
新华书店总店北京发行所经销
江西教育印刷厂印刷

*

850×1168 毫米 大 32 开 13.125 印张 370 千字

2001 年 4 月第 1 版 2001 年 4 月第 1 次印刷

定价:23.00 元

* * * *

(如有印装质量问题可更换)

编写说明

《概率论与数理统计》是作为财经与管理各专业本科生教材而写的.本书着眼于财经与管理科学的发展前景,为适应培养高水平财经与管理人才的需要,除了注重概率统计的基础知识之外,还纳入了有着广泛应用的各种统计方法,并简单介绍了非常有用的大型统计分析软件 SAS,以改变我国概率统计教学与计算机脱离的现状.本书也可作为自学、函授及夜大教材,还可作为研究生及经济工作者的参考书.

本书由梅国平、袁捷敏、毛小兵、胡长和编著.梅国平负责编写第0章、第一章、第七章,袁捷敏负责编写第二章、第八章、第十章,毛小兵负责编写第三章、第四章、第十一章,胡长和负责编写第五章、第六章、第九章.全书的写作大纲和统纂工作由梅国平教授负责.谭光兴教授、余仲弓教授、易伟明教授和陶长琪博士对本书的编写提出了不少宝贵意见.

对江西财经大学信息管理学院、江西财经大学教务处及南昌高校教育图书有限公司的支持和帮助表示感谢!

由于作者水平有限,加之时间仓促,书中错误在所难免,恳请读者不吝指正。

目 录

第 0 章 引论	(1)
0.1 概率论与数理统计发展简史	(1)
0.1.1 概率论发展简史	(1)
0.1.2 数理统计学发展简史	(5)
0.2 数理统计学的基本内容与应用	(12)
0.2.1 数理统计学的基本内容	(12)
0.2.2 数理统计学的应用	(15)
第一章 随机事件及其概率	(19)
1.1 随机事件	(19)
1.1.1 随机试验	(19)
1.1.2 样本空间	(20)
1.1.3 随机事件及其发生	(20)
1.1.4 事件之间的关系和运算	(21)
练习 1.1	(26)
1.2 随机事件的概率	(27)
1.2.1 概率的统计意义	(27)
1.2.2 概率的古典定义	(28)
1.2.3 概率的公理化定义	(34)
1.2.4 概率的性质	(34)
练习 1.2	(37)
1.3 条件概率与事件的独立性	(38)
1.3.1 条件概率	(38)
1.3.2 事件的独立性	(43)
1.3.3 独立试验及贝努里试验模型	(47)
练习 1.3	(49)
1.4 全概率公式与逆概率公式	(49)
1.4.1 全概率公式	(50)

1.4.2 逆概率公式(Bayes 公式)	(51)
练习 1.4	(53)
习题一	(54)
第二章 随机变量及其分布	(57)
2.1 随机变量及其概率分布	(57)
2.1.1 随机变量	(57)
2.1.2 随机变量的分布函数	(59)
练习 2.1	(61)
2.2 离散型随机变量及其分布律	(61)
2.2.1 离散型随机变量的分布律	(61)
2.2.2 离散型随机变量的分布函数	(63)
练习 2.2	(64)
2.3 几种常见的离散型分布	(66)
2.3.1 两点分布	(66)
2.3.2 二项分布	(66)
2.3.3 泊松(Poisson)分布	(69)
2.3.4 超几何分布	(72)
练习 2.3	(74)
2.4 连续型随机变量及其密度函数	(74)
2.4.1 密度函数	(74)
2.4.2 有关事件的概率	(75)
2.4.3 几种常见的连续型分布	(77)
练习 2.4	(80)
2.5 正态分布	(81)
2.5.1 正态分布的密度函数及其特点	(81)
2.5.2 标准正态分布	(82)
2.5.3 一般正态分布与标准正态分布的关系	(86)
练习 2.5	(87)
2.6 随机变量的函数及其分布	(88)
2.6.1 随机变量函数的定义	(88)

2.6.2 离散型随机变量函数的分布·····	(88)
2.6.3 连续型随机变量函数的分布·····	(90)
练习 2.6 ·····	(92)
习题二 ·····	(92)
第三章 多维随机向量及其概率分布 ·····	(97)
3.1 随机向量及其联合分布函数·····	(97)
3.1.1 多维随机向量·····	(97)
3.1.2 联合分布函数的性质·····	(98)
3.1.3 边缘分布函数·····	(99)
3.2 二维离散型和连续型随机向量·····	(99)
3.2.1 二维离散型随机向量 ·····	(100)
3.2.2 二维连续型随机向量 ·····	(105)
练习 3.2 ·····	(114)
3.3 随机变量的独立性 ·····	(115)
练习 3.3 ·····	(120)
3.4 随机向量的函数及其概率分布 ·····	(121)
3.4.1 和的分布 ·····	(122)
3.4.2 最大值与最小值的分布 ·····	(126)
3.4.3 线性和的分布 ·····	(129)
3.4.4 差、积、商的分布 ·····	(131)
练习 3.4 ·····	(136)
习题三 ·····	(136)
第四章 随机变量(向量)的数字特征 ·····	(140)
4.1 随机变量的数学期望 ·····	(140)
4.1.1 离散型随机变量的数学期望 ·····	(140)
4.1.2 连续型随机变量的数学期望 ·····	(144)
4.1.3 随机变量函数的数学期望 ·····	(146)
4.1.4 数学期望的性质 ·····	(148)
练习 4.1 ·····	(151)
4.2 随机变量的方差 ·····	(152)

4.2.1	方差的概念	(152)
4.2.2	方差的性质	(154)
4.2.3	矩	(158)
练习 4.2		(159)
4.3	随机向量的数字特征	(159)
4.3.1	协方差和相关系数	(159)
4.3.2	随机向量的数学期望向量,协方差矩阵与相关矩阵	(163)
4.3.3	n 维正态分布	(165)
练习 4.3		(167)
习题四		(167)
第五章	大数定律与中心极限定理	(171)
5.1	大数定律	(171)
5.1.1	依概率收敛的概念	(171)
5.1.2	切比雪夫不等式	(172)
5.1.3	切比雪夫大数定律	(175)
5.1.4	贝努里大数定律	(176)
练习 5.1		(178)
5.2	中心极限定理	(178)
练习 5.2		(184)
习题五		(185)
第六章	数理统计的基本概念	(188)
6.1	总体与样本	(188)
6.1.1	总体与个体	(188)
6.1.2	样本和简单随机样本	(189)
6.1.3	统计量	(189)
练习 6.1		(191)
6.2	数理统计中的某些常用分布	(192)
6.2.1	χ^2 分布	(193)
6.2.2	t 分布(“学生”氏分布)	(194)

6.2.3 F-分布	(196)
练习 6.2	(199)
6.3 抽样分布	(199)
6.3.1 一个正态总体的抽样分布	(199)
6.3.2 二个正态总体的抽样分布	(202)
练习 6.3	(206)
习题六	(207)
第七章 参数估计	(210)
7.1 参数的点估计	(210)
7.1.1 矩估计法	(211)
7.1.2 最大似然估计法	(213)
练习 7.1	(220)
7.2 点估计量的评价标准	(221)
练习 7.2	(226)
7.3 参数的区间估计	(227)
7.3.1 单个正态总体参数的区间估计	(230)
7.3.2 两个正态总体均值差和方差比的区间估计 ..	(237)
练习 7.3	(242)
7.4 比率的区间估计	(245)
7.4.1 比率	(245)
7.4.2 比率的区间估计	(246)
练习 7.4	(249)
习题七	(250)
第八章 假设检验	(253)
8.1 假设检验的基本概念	(253)
8.1.1 假设检验的概念	(253)
8.1.2 假设检验的基本原理	(254)
8.1.3 统计假设检验可能犯的两类错误	(254)
8.1.4 统计假设检验的一般步骤	(256)
8.1.5 参数假设检验与区间估计之间的关系	(259)

练习 8.1	(259)
8.2 单个正态总体的参数假设检验	(260)
8.2.1 方差已知的正态总体的均值检验	(260)
8.2.2 方差未知的正态总体的均值检验	(265)
8.2.3 单个正态总体的方差检验	(268)
8.2.4 在大样本场合下,非正态总体的均值检验	(272)
练习 8.2	(273)
8.3 两个正态总体的参数假设检验	(274)
8.3.1 已知 σ_1^2, σ_2^2 时,两个正态总体的均值检验	(274)
8.3.2 未知方差 σ_1^2, σ_2^2 (但 $\sigma_1^2 = \sigma_2^2$) 两个正态总体的均值检验	(275)
8.3.3 两个正态总体的方差检验	(278)
练习 8.3	(284)
8.4 总体比率的假设检验	(285)
8.4.1 单总体的比率检验	(285)
8.4.2 两个总体的比率检验	(289)
练习 8.4	(291)
8.5 非参数假设检验	(292)
8.5.1 拟合优度检验	(292)
8.5.2 正态性检验介绍	(294)
练习 8.5	(295)
习题八	(295)
第九章 方差分析	(302)
9.1 单因素方差分析	(302)
练习 9.1	(313)
9.2 二因素方差分析	(313)
9.2.1 无交互作用的方差分析	(314)
9.2.2 有交互作用的方差分析	(319)
练习 9.2	(327)
习题九	(327)

第十章 回归分析	(330)
10.1 回归概念	(330)
10.1.1 变量之间关系的两种类型	(330)
10.1.2 相关关系	(330)
10.1.3 回归分析	(331)
练习 10.1	(332)
10.2 一元线性回归分析	(332)
10.2.1 一元线性回归模型	(332)
10.2.2 样本线性回归、方程及其建立	(333)
10.2.3 线性样关关系的显著性检验	(337)
10.2.4 预测	(342)
10.2.5 控制	(344)
练习 10.2	(345)
10.3 可线性化的一元非线性回归	(345)
10.3.1 双曲线模型	(346)
10.3.2 半对数曲线模型	(346)
10.3.3 倒指数曲线模型	(346)
10.3.4 S 型曲线模型	(347)
练习 10.3	(348)
10.4 多元线性回归分析简介	(349)
10.4.1 多元线性回归的数学模型	(349)
10.4.2 样本线性回归方程的建立	(350)
10.4.3 线性相关关系的显著性检验	(351)
练习 10.4	(354)
习题十	(354)
第十一章 SAS 软件简介	(357)
11.1 SAS 的安装、启动与退出	(357)
11.2 SAS/INSIGHT	(357)
附录表	(403)
参考文献	(377)

*第0章 引 论

概率论和数理统计都是应用数学的重要分支,但就它们二者之间的关系来说,通常认为概率论是数理统计的基础,数理统计则是概率论的一种应用.在接触本学科的基本理论与方法之前,了解其发展简史、特点以及应用情况,有助于对本学科的深入领会与实际运用.这就是本章要讲述的内容.

0.1 概率论与数理统计发展简史

0.1.1 概率论发展简史

我们将概率论的发展历史大致上划分为以下五个阶段.

1. 概率论发展的原始阶段

这一阶段很难确定它发展的起端,有些学者把它的发展追溯到人类最早期的文明史.在史前的圣堂和神殿中曾经发现过一些用马的腿骨制成的小立方体,这表明在早期的宗教仪式中可能使用过骰子.至15世纪,在意大利兴起了航运保险事业,这说明那时人们至少已有了概率论的初步知识.

2. 概率论发展的早期阶段

这一阶段大体上从17世纪中叶开始,至18世纪初期结束.主要的代表人物是法国的巴斯加尔、费尔马和荷兰的惠更斯.突出的事件是概率论在赌博问题中的应用.1654年,一些热衷于赌博的人向巴斯加尔提出了还不能归入那时数学范畴的划分赌本的问题,这引起了巴斯加尔的兴趣.他把解法写信告诉了费尔马,从而二人采用通信的方式展开了讨论.现在这些信件被公认为是概率论中最早有详尽文字记载的文献.基于这一原因,人们通常把这看

成是概率论发展的起点. 在费尔马和巴斯加尔的讨论中, 已提炼出了像概率和数学期望这样一些概率论中最重要的概念. 稍后, 另一位年轻的荷兰人惠更斯也独立地讨论了类似的赌博问题, 并于 1655 年发表了概率论中的第一篇论文《赌博中的计算》. 这一阶段, 由于科学发展水平不高, 所以充满机遇性的赌博便成为那时发展概率论的直观背景. 解决概率论中问题时用到的数学工具也仅是初等的排列和组合算法. 这里我们需要强调一下, 尽管概率论的早期应用主要限于赌博, 但在讨论赌博问题中发展起来的思想却对于认识随机现象有着普遍的意义. 上面提及的几位数学家在当时就已觉察到了这点. 惠更斯在《赌博中的计算》一文中写到: “我相信如果读者仔细地研究这些材料, 当可发现你所处理的不仅仅是赌博, 其中实际上包含了既深刻又有趣的一种新理论的基础.” 概率论以后发展的历史确实验证了惠更斯的预言.

3. 概率论发展的中期阶段

这一阶段从 18 世纪初期开始, 一直延续到 19 世纪中叶. 这一阶段中, 概率论的应用已不局限于赌博, 而是扩展到观察误差理论、射击学、人口统计学等众多领域. 解析的方法也被引入到概率论的研究之中. 这一时期的代表人物是以拉普拉斯为首的一批法国数学家, 其中比较杰出的还有贝努里、德莫哇佛尔和泊松等人. 此外, 英国的贝叶斯和德国著名数学家高斯对于概率论发展作出的贡献也不容忽视.

贝努里的遗作《猜测的艺术》于 1713 年发表, 这一事件被认为是概率论这一发展阶段的开端. 现在称之为贝努里大数定律的概率论中的第一个极限定理便是首次在此书中刊出的.

拉普拉斯领导了概率论在这一发展阶段中的潮流. 他一生发表了不少概率论方面的论文, 其中影响最为深远的是于 1812 年发表的集大成作品《分析概率论》. 在这一著作中, 他不仅总结了自己的研究成果, 而且还系统地整理了概率论至那时为止已取得的主要进展. 他还利用等可能性事件的概念明确地给出了概率的古典定义, 从而建立了古典概率论基础的理论框架. 直到 20 世纪初, 关

于概率论的教科书还一直沿用他所给出的概率论的古典基础,由此足见他这部著作对后世所产生的影响.拉普拉斯不仅研究一般的理论问题,而且热衷于概率论的应用.他成功地运用概率论去估计观察误差,在人口统计学的研究中也有所建树.

贝叶斯的重要论文《论机遇理论中一个问题的解决》是他逝世后两年,即 1763 年发表的.这篇论文中包含了归纳推理的一种思想,在当时并未引起人们足够的重视,现在已被发展成为关于统计推断的一种系统的理论和方法.

高斯系统地建立了观察的误差理论,推导了随机误差的分布规律,即正态分布律,这是概率和统计中最重要的分布律之一.他在研究观察数据的随机误差时所提出的最小二乘法,也在以后统计学中发展起来的回归分析中找到了重要的应用.

这一阶段,概率论在理论中取得的另一进展是,自从贝努里开始,德莫哇佛尔、拉普拉斯和泊松等人都研究了概率论中古典的极限定理.这些定理既有助于深刻理解概率论的基础,而且诱发了以后更为深入的工作,使得极限定理成为概率论研究中久盛不衰的课题.

综上所述,这一阶段中法国古典概率学派处于概率论发展的中心地位,在他们的努力下,古典概率论得到了全面的发展,在应用方面也取得了长足的进步.但是这种发展势头到了 19 世纪后半世纪却在法国出现了停滞,这主要归因于古典概率基础的局限性.这样概率论发展的中心移出了法国,开始了一个新的发展时期.

4. 概率论发展的近期阶段

这一阶段经历的时间大体上可认为是 19 世纪后半世纪以及 20 世纪的前 40 年.这一阶段中,概率论的应用范围不仅继续得以扩展,而且也更为深入.概率论应用的一个重要方面是统计学.以皮尔逊和费希尔为首的英国统计学派领导了统计学这一时期的发展潮流.现代数理统计学中几乎所有重要的统计理论和统计方法都是他们在 19 世纪末和 20 世纪初这一段时间内建立起来的,

统计学从而成为应用数学的一个独立的分支.这方面的背景材料我们将在稍后详细地介绍.

概率论在这一阶段中令人瞩目的另一重要应用是在统计物理学方面.对此作出贡献的代表人物是奥地利物理学家玻茨曼和美国物理学家吉布斯.他们首先将概率论和统计学的论证方法贯穿到分子物理学的研究中.鉴于大量分子的存在以及它们热运动的极其紊乱的性质,使得概率论和统计学在这一领域中的应用获得了极大的成功.概率论在这样一种严肃科学的应用中取得了成功,而且这还是由物理学家实现的,就不得不使人们对概率论刮目相看了.概率论在物理中的应用至今仍是概率论研究中相当活跃的领域之一.

这一阶段在概率论的理论研究工作中取得突破性成就的,当首推俄罗斯和前苏联概率论学派.他们中间的代表人物是切比雪夫、马尔科夫、伯恩斯坦、欣钦和柯尔莫哥洛夫等人.这一阶段中他们对于概率论作出的主要贡献可以归纳为两个方面.其一是继续在由贝努里和拉普拉斯开创的古典极限定理的研究方向上进行探索,但他们已把所研究的内容大大地推广了.

前苏联概率论学派作出的另一重大贡献是彻底改造了概率论的古典基础.随着时间的推移,概率论古典的局限性就越来暴露出来了.它极大地限制了概率论在物理、生物、统计等领域中的进一步应用.这样到了 20 世纪初,对于概率论的古典基础进行改造已是一种客观上的要求.这时的数学界正风行公理化的潮流,由罗巴契夫斯基在研究非欧几何时所建立起来的公理化方法在那时已渗透到数学的各个分支,于是一些卓越的数学家自然会想到以公理化手段来重建概率论的逻辑基础.概率论公理化的先行者是伯恩斯坦,他在 1917 年发表了第一篇关于概率论公理基础的随笔.概率论公理化的最终完成是由柯尔莫哥洛夫实现的.他在 1933 年发表的《概率论的基本概念》一书中,系统地表述了现在已被广为接受的概率论公理系统,第一次把概率论建立在严密的逻辑基础之上.从此概率论确立了它作为一门严格数学分支的地位,它的发展也开始了一个全新的阶段.

5. 概率论发展的现期阶段

这一阶段可认为是从第二次世界大战结束后直至现在的一段时期. 在这一段时间内, 概率论发展的步伐极大地加速了, 其理论研究工作的内容得以大幅度地拓广, 深刻程度也远非古典概率论所能比拟. 概率论在应用方面也取得了辉煌的成就, 许多应用概率论的分支相继形成.

0.1.2 数理统计学发展简史

我们分三个阶段来介绍统计学的发展历史.

1. 统计学发展的初级阶段

我们把 19 世纪中叶以前都划归为这一阶段. 统计学在这一阶段的发展过程中反映出以下二种变化趋势:

- (1) 由纯粹收集数据转变到借助数据进行简单的推断;
- (2) 由最初和概率论互相分离到逐渐结合.

先谈第一种变化. 统计学在英语中称为 Statistics. Statistics 源出于拉丁文中的“Status”一词, 它和英语中的“State”(国家)同义. 所以在西方, 统计最早指的就是收集国情资料. 若从这个意义上去理解统计学的发展, 它的历史就源远流长了, 和概率论发展的情形一样, 几乎也可追溯到人类的早期文明史. 拿我国来说, 早在公元前 2000 多年前的周朝就保存有收征钱粮和朝廷消费支出的大量记录, 也保存有人口和征兵的记录. 但类似这种纯粹属于收集数据的统计工作, 和我们现在对于统计学的理解相去甚远. 这种情况一直延续到 17 世纪中叶才有所变化.

西方学术界普遍认为统计学的研究事实上是由英国人格兰特开始的, 并称他为人口统计学之父. 人口统计学是从事人类总体的统计研究的. 格兰特是一个商人, 他在经营服饰及缝纫用品的同时, 潜心研究死亡记录, 统计了不同死因的百分比. 结果他发现死于不同原因, 如自杀, 意外事故和某种疾病的百分比, 不仅极有规律, 而且逐年之间几乎保持不变. 格兰特通过统计分析还发现, 男

婴的出生要比女婴多,但由于男人因职业上公害(工业废气和废水对健康带来的损害)、疾病和战争致死的较多,所以到了结婚的年龄,男人和女人的数目近乎相等.这些观察结果都总结在他于1662年发表的《基于死亡记录的自然和政治方面的观察》一书中.由于他在人口统计学上作出的贡献,他成为于1660年成立的英国最古老的科学协会——皇家学会的特约会员.格兰特的上述著作作为以后死亡率表的制定铺平了道路,也对统计学上一些先驱者的工作产生了影响,其中包括预言哈雷彗星回归的天文学家哈雷.在1700年,哈雷发表了他关于死亡率表和寿命估计的研究成果.这些成果对于人寿保险事业当然是很有用的,这使人们感到,统计推断作为一门科学而言,不仅可信的,而且是值得花费时间和精力.

从格兰特和哈雷的工作开始,统计学的发展已具有这样一种特色:针对某些特定的问题进行观察试验,收集资料,且已不满足于数据的简单统计,有时还能越出收集数据的范围,进行某种简单的推断.这比人类早期纯属例行公务式地收集数据的活动已有不少长进.但必须指出,这样一种推断只是基于一种朴素直观的想法,还没有坚实的理论作为其后盾.出现这种情况也是很自然的,因为作为现代数理统计基础的概率论的发展,在那时也才刚刚起步.这样,直至贝努里《猜测的艺术》一书的发表,尽管概率论和统计学均已被人们承认,但它们的发展基本上是互相平行的,其内容也互相分离.这一情况在随后的一个半世纪中才有所变化,概率论的思考方法开始逐步进入统计学的研究之中.这就是我们要谈的统计学在这一阶段中发生的第二种变化.

前面已提及,从18世纪初开始至19世纪中叶是古典概率论全面发展的时期.在这一时期内,由格兰特开创的人口统计学研究成为概率论的重要应用领域之一.古典概率论的奠基人拉普拉斯曾对此产生过浓厚的兴趣.他在根据一次讲演写成的《概率的哲学探讨》一文中,讨论了许多人口统计学方面的具体问题.他运用概率论中的概念叙述了死亡率表的制定方法,引入了平均寿命的概念.特别地,他还考虑了如何根据死亡率表去计算人口总数的问题.