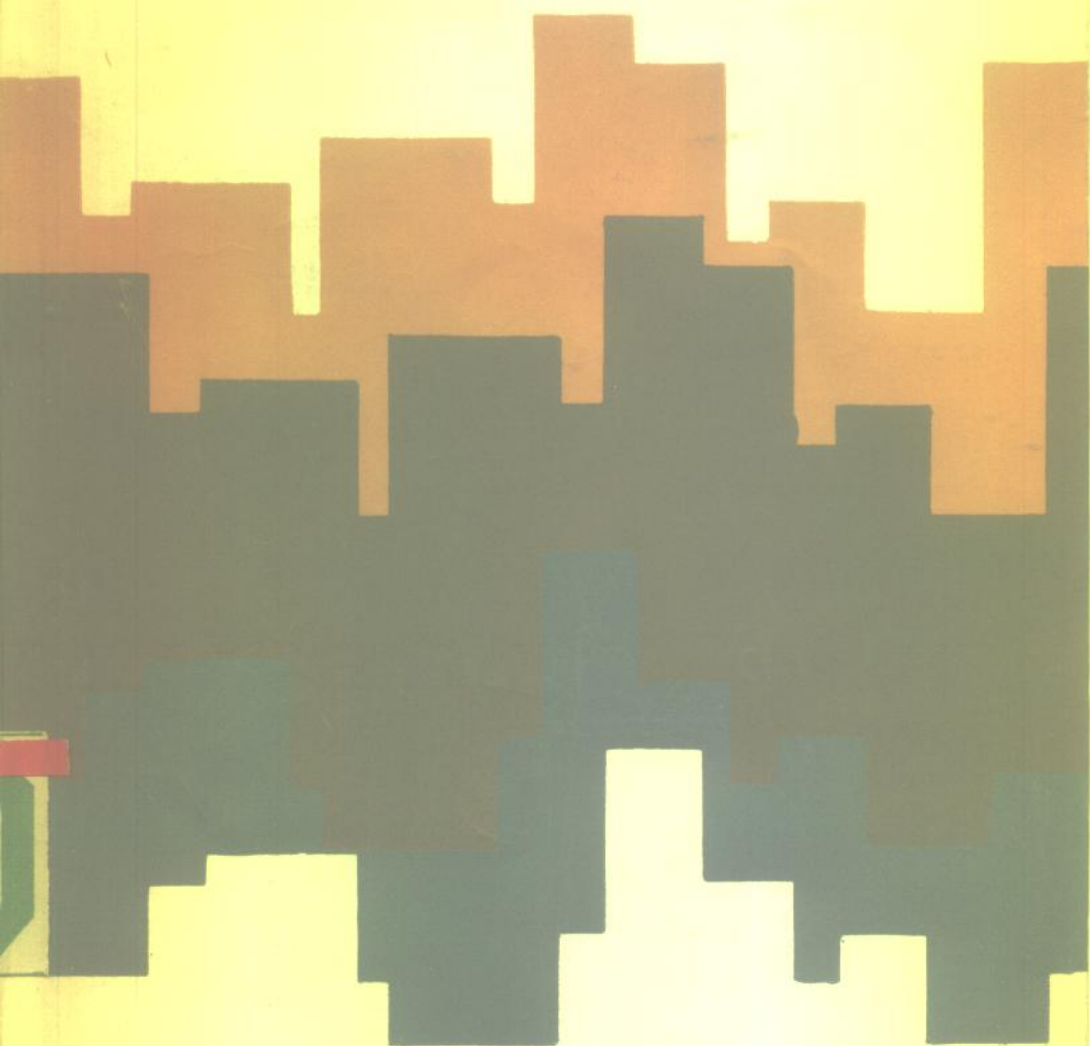


社会统计学

〔美〕布莱洛克 著



社会统计学

〔美〕布莱洛克 著

傅正元 沈崇麟 黎 鸣 译
赵 平 叶念先 章扬恕

斜 川 审校

中国社会科学出版社

1988·北京

责任编辑：徐菊芳
封面设计：式 一
版式设计：钱 锋
责任校对：徐培英

社会统计学

SHEHUI TONGJIXUE

〔美〕布莱洛克著

傅正元等译

*

中国社会科学出版社 出版
发 行

新华书店 经销

北京新华印刷厂 印刷

850×1168毫米 32开本 20.25印张 2插图 500千字

1988年12月第1版 1988年12月第1次印刷

印数 1—5 000册

ISBN 7·5004·0269·4/Z·58 定价：6.30元

译者前言

自本世纪四十年代以来统计学逐步广泛地应用于生物科学和社会科学各学科。在西方, 统计学 statistics 一词在 1850 年之前尚被人们理解为与国家 state 事物有关的数字和报表。但此后由于数理统计学家和应用统计学家的努力, 人们逐渐理解统计学的作用不仅在于用数字描述客观情况, 而且在于从数据资料中总结规律, 推论更大范围的情况, 做出判断和帮助设计实验和调查。社会科学的发展有赖于实际调查, 社会统计学正是为了帮助社会科学工作者设计调查方案, 分析数据资料、做出结论而发展起来的。

目前社会科学各学科如社会学、政治学、数量经济学、法律学、管理学及生物科学, 如心理学、医学、生物学等等几乎都将统计学作为学科的主要研究和调查方法的基础。社会统计学已成为世界上各高等院校的社会科学系科的必修课。

本书作者布莱洛克教授是国际知名的社会统计学家和社会学家,《社会统计学》一书自 1960 年出版以来一直是美国大学中被广泛采用的教材, 同时也流传到欧洲各国高等院校, 并被多次重印。本译本根据 1979 年修订本翻译。

统计学在我国社会科学界尚未广泛使用, 但这正是使我国社会科学现代化的一个主要方面。社会统计学的应用并不需要高深的数学知识, 阅读和学习本书的读者仅需知道中学代数即可, 可是要求读者学习好统计逻辑思维的方法。本书用通俗的语言介绍读者了解和学会使用一些实际可行的统计方法。

由于一些术语在我国尚无统一译法，如“tie”、“probit”、“odds”、“logit”等仅能根据我们的理解暂定译名，书后附有英汉社会统计学术语对照表。

希望本书有助于读者了解社会统计学的方法和内容。

我国翻译界在译者前言、后记中一般都用下面几句话作为结尾，我们也不例外，由于译者水平不高，译文中不当之处请大家帮助指正。

傅正元

一九八五年三月

GDZ 81/13

前 言

本书主要供将要从事社会研究的社会学高年级大学生和研究生阅读。

本书初版问世以来的十九年中，应用统计学的教学水平和方法理论都有很大发展，不仅在社会学方面是这样，而且在政治学、社会人类学、地理学及社会工作等各方面也是这样。然而，这些学科的绝大多数学生和研究人员还缺乏必要的数学知识，无法充分利用数理统计和计量经济学日益增多的专业文献。为此，本书尽可能避免数学推导，对一般学生来说，只要复习一下附录1的代数知识就可以阅读本书。虽然在初级统计学教程中不必注重数学推导，但作者认为，如果要得到较为完整、较为系统的统计学知识，而不是知其然不知其所以然，仍必须透彻理解统计推论原理的某些基础的、带根本性的思想。所以，本书比较注意统计推论的基本逻辑，还用专门的一章讲概率论，而对一般初级教程中按惯例讲述的一些题目则不予过多注意。

在应用统计学的教学中最困难的问题之一是如何提高学生的兴趣和主动性。第一，要使他们克服对数学的畏惧心理，第二，要让他们学会把统计学应用到各自的专业上去。为了达到第二个目的，作者没有涉及太广的应用范围，而是选择一些社会学家普遍感觉兴趣的实例。还有些例子选自与社会学相近的学科和领域，如社会心理学、社会工作和政治行为。在多数情况下，每一

新的课题用只一个实例说明，因为实例过多很可能使学生不得要领。更多的实例将作为习题安排在各章的末尾。一般说来，作者尽可能简单明了地阐述基本原理，同时，在讨论每一新的题目时对一些较难理解的概念予以必要的重复。介绍新概念时尽可能做到循序渐进，同样，讲述新内容时力求与先前讲过的内容联系起来。这样做的主要目的是使学生看出，许多最常用的检验方法和量度方法基本上是相似的。

希望改进本书第一版和第二版的读者几乎都建议增加内容，他们还建议，许多原已论及的题目应给予专业性更强的介绍。我个人认为，社会学家和政治学家确实需要较多地接触专业性更强的文献，如关于实验设计和在非实验性研究中使用联立方程的材料。但显然，如果增加这些内容，本书就不再适宜给主修社会科学的高年级大学生充当入门教材。

所以最后决定本书不讨论以下问题：计算机程序设计、一般线性模型、实验设计、联立方程程序、路径分析以及量度误差的处理方法等，这些问题在我与过去的两位同事——刘易斯·卡特和克里施南·南布狄里——所合写的题为“实用多元分析与实验设计”一书中均有阐述，该书已由麦格劳-希尔公司出版，可作为本书的补充读物。

现在这个修订版，除在每章末尾增加了习题和新近出版的书目之外，几乎所有其他改动和补充都有一个目的：向学生介绍一些更高深更专门的题目，这些题目对社会学家来说越来越有价值。除少数例外，这些补充都在第15章到第20章里面，而且大部分都涉及多元分析。

在讲述定类尺度程序的第15章，以长短不一的篇幅补充进以下各方面的材料：似然比率卡方、卡方分解、可能性比率、预测逻辑及统计交互作用。18、19两章对以下问题的讨论较前更为详尽：关系的定序量度、分组注意事项和量度误差、奎德定序量

度的偏相关匹配处理程序、定序量度的显著性检验。扩充关于定类和定序程序的内容，目的是帮助学生更轻松地阅读有关这方面的专业文献，现在这类程序比以前容易搞到，而且已经成为标准计算机程序包(如 SPSS)的组成部分。

现在第 16 章的开头简短讨论一般线性模型，此外，在该章和第四部分其他各章还会着重讨论一般线性模型。实验设计的内容没有增加，因为这个问题在本书的姐妹篇《实用多元分析与实验设计》一书中足足占了三分之一的篇幅。第 19 章增加了一点因果模型和路径分析的内容，并简单介绍一般线性模型的矩阵代数表达方法，尽管这些问题在《实用多元分析与实验设计》一书中有详尽的介绍。第 20 章增加了对桑奎斯特和摩根的自动交互检测程序和对数线性模型的阐述。

学好统计学必须做数字习题，还必须熟练掌握计算机程序，所以由詹姆斯·亨内写了一本教师手册作为本教材的补充读物。

在亨内的教师手册中有一些习题，可以帮助理解本书的各个章节，还可以帮助理解 SPSS 手册，所以希望同时使用这两本书的学生和教师能从教师手册得到更多的指导。南布狄里、卡特和布莱洛克的教材(即《实用多元分析与实验设计》)中有一章讲述 FORTRAN 语言，使读者学会修改程序，以便解决非标准类型的统计问题，这类问题是 SPSS 一类的程序包中所没有的。

本书的某些章节和习题在概念上难于理解，或是要求学生已学过研究方法的课程，这些地方都用星号(*)标示，初次阅读可以略去。如果计划用一个学期教完本课程，这些内容可以不讲。

在本书写作和修改过程中，密执安大学、耶鲁大学、北卡罗莱纳大学、华盛顿大学的许多学生为改进本书提供了大量的建议，理查德·拉皮埃尔、桑福德·多恩布希、罗伯特·埃利斯、桑托·卡米勒里、西奥多·安德森、理查德·艾姆斯、埃里卡·博登、路易斯·古德曼曾读过本书初稿及所印各版并提出批评意

见，安·布拉洛克、黛安·埃策尔、安·劳克斯、多丽丝·斯莱辛杰曾为本书校对、打字、验算，在此谨致谢意。

我深深地感谢丹尼尔·普赖斯，因为正是他激发了我对统计学的兴趣。

原剑桥大学教授罗纳德·费歇尔爵士、弗兰克·耶茨博士以及爱丁堡的奥利弗与博伊德公司同意利用他们的《供生物学、农学和医学研究用统计图表》制成本书图表ⅢⅣⅤ，在此谨致谢意。我也同样感谢其他出版家和作者，他们允许我利用各种图表和计算表格。

小休伯特·布莱洛克

目 录

前 言	1
-----	---

第一部分 导论

第1章 导言：统计学的作用和局限性	3
1.1 统计学的作用	4
1.2 统计学在研究过程中的地位	7
1.3 一点忠告	8
第2章 理论、量度和数学	10
2.1 理论和假设：操作性定义	10
2.2 量度层次：定类尺度、定序尺度和 定距尺度	14
2.3 量度与统计量	19
2.4 本书的结构	23

第二部分 单变量描述性统计学

第3章 定类尺度：比例、百分数、比率	31
3.1 比例	31
3.2 百分数	33
3.3 比率	37
第4章 定距尺度：频数分布和图示法	41
4.1 频数分布：数据分组	41

4.2	累积频数分布	47
4.3	图示法: 直方图、频数多边形积和 累积频数分布曲线	48
第5章	定距尺度: 集中趋势的量度	54
5.1	算术平均值	55
5.2	中位数	58
5.3	从分组数据计算均值和中位数	60
5.4	均值与中位数比较	65
5.5	集中趋势的其他量度	68
5.6	十分位数、四分位数和百分位数	70
第6章	定距尺度: 离差的量度	73
6.1	极差	73
6.2	四分位偏差	75
6.3	平均偏差	75
6.4	标准差	76
6.5	变差系数	81
6.6	其他总结性量度	82
第7章	正态分布	84
7.1	有限的与无限的频数分布	84
7.2	正态曲线的一般形式	87
7.3	正态曲线下的面积	89
7.4	使用正态分布表的说明	92

第三部分 归纳性统计学

第8章	归纳性统计导论	99
8.1	统计量和参数	99
8.2	假设检验的步骤	100
8.3	以推论结果证实前提的谬误	103
8.4	统计性假设的形式	104

第 9 章 概 率	109
9.1 先验概率	110
9.2 概率的数学性质	114
9.3 排列	125
9.4 期望值和矩	132
9.5 独立性和随机抽样	134
第 10 章 假设检验: 二项分布	144
10.1 二项概率分布	144
10.2 统计检验的步骤	149
10.3 二项分布的应用	163
10.4 二项分布的推广	167
10.5 小结	169
第 11 章 含均值和比例的单样本检验	174
11.1 均值的抽样分布	174
11.2 已知 σ , 总体均值检验	182
11.3 学生 t 分布	186
11.4 关于比例的检验	191
第 12 章 点估计和区间估计	198
12.1 点估计	199
12.2 区间估计	202
12.3 其他类型问题的置信区间	208
12.4 确定样本容量	210

第四部分 双变量和多变量统计学

第 13 章 双样本检验: 均值差和比例差	219
13.1 均值差检验	220
13.2 比例差	227
13.3 置信区间	231
13.4 非独立样本: 配对	232
13.5 对实验设计和显著性检验的几点评论	234

第 14 章	定序尺度：双样本非参数检验	243
14.1	检验效力和检验效率	244
14.2	瓦尔德-沃尔夫维茨链检验	250
14.3	曼-威特尼或威尔柯克森检验	255
14.4	柯尔莫哥洛夫-斯米尔诺夫检验	263
14.5	威尔柯克森配对符号-秩检验	266
14.6	小结	270
第 15 章	定类尺度：列联问题	275
15.1	卡方检验	275
15.2	费舍尔精确检验	287
15.3	卡方分解和其他检验	292
15.4	关系强度的量度	294
15.5	其他变量的控制	311
15.6	有关分类变量的注意事项	321
第 16 章	方差分析	332
16.1	简单的方差分析	332
16.2	特定均值的比较	343
16.3	双向方差分析	348
16.4	代替方差分析的非参数方法	362
16.5	关系的量度：组内相关	367
第 17 章	相关和回归	376
17.1	线性回归和最小二乘方	377
17.2	相关	391
第 18 章	相关和回归(续)	409
18.1	显著性检验和置信区间	408
18.2	非线性相关和回归	420
18.3	量度误差的影响	426
18.4	定序尺度：秩序-相关	428
18.5	判定量度和分类层次	438

目
录

第 19 章	多重相关与偏相关	446
19.1	多重回归和最小二乘方	446
19.2	偏相关	449
19.3	偏秩-序相关	457
19.4	偏相关和因果解释	464
19.5	多重最小二乘方和 β 系数	473
19.6	多重相关	478
19.7	多重回归和非线性	484
19.8	显著性检验与置信区间	489
19.9	一般线性模型的矩阵代数表示法	493
第 20 章	协方差分析、虚构变量、 线性模型的其他应用	504
20.1	关联两个定距尺度, 控制定类尺度	505
20.2	关联定距和定类尺度, 控制定距尺度	521
20.3	协方差分析的推广	527
20.4	虚构变量分析	528
20.5	多重分类分析和自动交互作用检测分 析	532
20.6	分类应变量: 对数线性模型	535
20.7	小结	539

第五部分 抽 样

第 21 章	抽样法	549
21.1	简单随机抽样	550
21.2	系统抽样	555
21.3	分层抽样	557
21.4	整群抽样	565
21.5	非概率抽样	570
21.6	非抽样误差和样本容量	571

附录 1	代数运算复习	574
附录 2	表	582
附录 3	英汉社会统计学术语对照表	620

第一部分 导论

