

〔英〕罗德里克·弗拉德著

历史计量法导论

商 务 印 书 馆



历史计量法导论

〔英〕罗德里克·弗拉德 著

肖 朗 刘立阳 等译

庞卓恒 校

商 务 印 书 馆

1992年·北京

Roderick Floud
**AN INTRODUCTION TO QUANTITATIVE
METHODS FOR HISTORIANS**

Methuen & Co., New York, 1979

根据纽约梅休因公司 1979 年版译出

LÌSHÌ JÌLIANGFĀ DǎOLUN

历史计量法导论

〔英〕罗德里克·弗拉德 著

肖 朗 刘立阳 等译

庞卓恒 校

商务印书馆出版
(北京王府井大街 36 号 邮政编码 100710)

新华书店总店北京发行所发行

河北香河县第二印刷厂印刷

ISBN 7-100-00884-0/K·146

1992 年 6 月第 1 版

开本 850×1168 1/32

1992 年 6 月北京第 1 次印刷

字数 155 千

印数 1 400 册

印张 6 7/8

定价：3.00 元

译者前言

本书作者罗德里克·弗拉德 (Roderick Floud, 1942—) 是英国当代历史学家。1964 年毕业于牛津大学, 1971 年获博士学位。大学毕业后, 他一直从事历史学的教学和研究工作。1966 年至 1969 年, 他在伦敦大学担任历史学讲师; 1969 年至 1975 年, 在剑桥大学担任讲师, 兼任研究员; 1975 年起, 在伦敦大学伯克贝克学院担任现代史教授。除本书以外, 其主要著作有《计量经济史论文集》(1974 年)、《1850 年至 1914 年英国机床工业》(1976 年) 等。

弗拉德以研究计量历史学及其方法而著称。一般认为, 计量历史学是继计量经济学之后, 把数理统计学的原理和方法运用于社会历史科学领域的新的尝试, 它发端于本世纪 50 年代末、60 年代初的美国。1958 年, 经济学家 A. 康拉德和 J. 迈耶尔率先运用计量方法研究经济史; 1961 年, 历史学家 L. 本森又把计量方法引入政治史研究。此后, 西欧一些国家的历史学家和社会科学家纷纷响应。在英国, 计量经济史得到迅速发展, A. 查普曼, W. G. 霍斯金斯以及本书作者被公认为代表著作家; 而法国历史学家则主要运用计量方法来研究社会史, 年鉴学派的主将 F. 布罗代尔, 勒·鲁瓦·拉杜里等人作出了卓越的贡献。随着计量历史学研究的广泛开展, 出现了一批介绍历史计量方法的论著, 本书就是其中的一部代表作。它不仅概括了计量历史学的基本方法, 而且凝聚着作者本人从事计量历史研究的心得, 因此, 1973 年出版后就得到学术界的好评, 遂于 1979 年再版, 并于 1983 年重印; 中译本就是根据这个本子翻译的。

按其内容，本书大致可分为三个部分。第一部分介绍了运用表、图和某些简单的数理统计方法对历史资料加以分类、整理和初步分析，着重叙述了频率分布、平均值、中位数、众数等计量集中趋势和离差的方法。第二部分探讨了与历史学密切相关的时间序列、变量相关以及证据短缺等问题，着重介绍了简单线性回归和随机抽样的概念及方法。第三部分描述了进行计量分析的工具，即电子计算机和电子计算器。在介绍这些内容时，作者引用了英国经济史、商业史和政治史中的许多资料，将它们作为实例一一加以示范性的分析、计算和检验。同时，考虑到许多史学工作者缺乏统计学和数学知识，凡涉及统计概念和数学计算，作者往往不厌其烦地作了详尽的阐述，由浅入深，循序渐进，以便使那些只具备初等数学知识的读者也能比较容易地把握书中的主要内容。鉴于这些特点，本书堪称为历史计量方法的入门书。

必须指出，数理统计方法只有在正确的观点指导下加以运用，才能得出正确的结论。对此，马克思主义经典作家早有论述。列宁在《俄国资本主义的发展》中写道：“把工人和业主加在一起作为一个整体，然后得出收支的‘平均数’，……这种平均数只是掩盖了下等户农民群众……的赤贫真相”^①。“社会认识的最有力的武器之一的社会经济统计，这样就变成了一种畸形的东西，变成了为统计而统计，变成了儿戏……”^②。本书作者也多次强调正确的历史理论对运用计量方法所具有的重要意义，并举例说明构造特定的理论模型是如何将具体的历史问题与史料证据相联系的。

本书正文由肖朗、刘立阳翻译，第九章及其附录由王薇翻译。全书由庞卓恒教授通校。杨秀珍协助了通校工作。

① 《列宁选集》，人民出版社 1963 年版，第 3 卷，第 125 页。

② 同上书，第 16 卷，第 431 页。

译者水平有限，译文不妥之处在所难免，恳盼读者提出宝贵意见。

1987年12月

目 录

导言.....	1
第一章 历史数据的分类.....	7
(1) 列名分类	8
(2) 顺序分类	10
(3) 差别或比例分类	11
(4) 某些复杂情况	12
(5) 重新分类与组合	15
第二章 历史数据的整理.....	17
(1) 数据集合	17
(2) 事件	18
(3) 变量	19
(4) 数据矩阵	19
(5) 数据的收集	23
第三章 简单的数学知识.....	28
(1) 频率分布	28
(2) 求和记号	34
(3) 对数	37
第四章 数据的初步分析 I: 频率分布及其图表.....	42
(1) 频率分布	44
(2) 交叉分组	48
(3) 图	50
(4) 比率标度图	58
第五章 数据的初步分析 II: 概要计量方法	66

1	(1) 算术平均值	66
	(2) 标准差	71
	(3) 几何平均值	75
	(4) 中位数	76
	(5) 众数	79
	(6) 变差系数	80
	(7) 运用哪一种概要计量方法?	80
第六章 时间序列分析		86
	(1) 时间序列分析的对象和假设	88
	(2) 增长率	91
	(3) 趋势	95
	(4) 时间序列的规律性波动	106
	(5) 比率和指数的用途	118
第七章 变量之间的关系		126
	(1) 是否存在一种关系?	127
	(2) 这种关系有多强?	138
	(3) 变量关系的形式	141
	(4) 时间序列数据的相关和回归	151
第八章 数据不完善问题		157
	(1) 信息过多,变量的选择	158
	(2) 信息过多,事件的选择	162
	(3) 抽样结果的“有效性”	173
	(4) 数据过少:数据短缺问题	176
	(5) 一个或多个事件数据短缺	177
	(6) 一个或多个变量数据短缺	179
	(7) 一个或多个事件的一个或多个变量的数据短缺,但不是整个 事件或所有变量的数据都短缺	183
第九章 计算器、计算机与历史数据		185
	(1) 设备的选择:电子计算器	186

(2) 设备的选择:计算机.....	188
(3) 为计算机准备历史数据	189
(4) 用计算机分析历史数据	190
参考书目	194
4位数对数表和反对数表	199
译名对照表	203

导 言

当我们描述和分析存在于过去或现在的人类社会时，不可避免地要使用数和量的概念。假如要对某个人作一番恰当而充分的描述，那么，他的年龄、出生日期、财产、娶过几个妻子、有多少子女——所有这些都是我们必须揭示的此人的数量特征。在这样做的时候，我们计量他的这些数量特征，并把他和另一些人作比较：是富一些呢还是穷一些，是年龄大一些呢还是小一些；通过这些方式，并通过对他的思想和业绩的讨论，以求把他放到他所生活的那个社会中去探讨。为了做到这一点，我们通常是把具有相似的行为方式或相似的思想的人分为一些群体。我们运用诸如“中产阶级”、“法国人”、“保守派”这样的术语去描述那些群体。我们必须这样来分门别类，因为只有通过这样的方法，才能把多种多样的人们的思想和行为归入一种易于处理的形式。

诸如年龄、财产、孩子数目这些度量显而易见都是定量性的。我们只有通过计算一个人出生后的年数，才能计量他的年龄；我们也只有通过用一定的货币单位的量来计算一个人所拥有的物质财富的数字或价值，才能计量他的财产。假如我们在描述生活于过去时代的人们时用了这样的度量，那么，我们就正是在用计量法。我们在历史著作中所用的另一些度量或描述方式，在形式上却是非定量性的，而且被用来描述个人或群体的思想或态度；“法西斯分子”、“文艺复兴时代的人”就属于这样的描述方式。但是，当我们在用诸如此类并非定量的、而是定性的描述时，常会发现只有通过计量持有这类观点或能被这类术语来描述的人们的数目，才

能赋予这些描述以充分的含义并评价它们的历史意义。“中产阶级”是对社会中一个集团的描述，但在许多情况下，它也是对社会中拥有特定收入和持有特殊态度的一定数目的人的描述。假如我们说：“中产阶级支持政府”，那就意味着即使不是其全部，也是我们描述为“中产阶级”的绝大多数人都支持政府，并且只有通过计算这类人的数目，才能最终确定这样一种陈述是否正确。因此，历史学家所运用的许多定性的判断或描述都隐含着一定定量的意义，有时候必须使这层含义明朗化。另外，对个人或集团的行为的许多描述都具有数量上的含义，如“通常”、“正常”、“经常”、“许多”这类字眼都与数量概念有关，尽管我们通常可能不想去准确地考查它们，但在原则上，它们的意义或正确性只有通过定量的计量才能被确立。

因此，历史学家就象其他社会科学家一样，都得经常地和不可避免地运用数量概念。这一事实并不意味着他们所作的全部陈述都是定量的，也不意味着他们相信人类行为的各个方面都能被计量并给出数字。人类、人类集团和物质对象的许多方面都不能用定量方式来估量或表述。尽管我们能够计算出一个中世纪农民的收入，我们却无法计量他在享用这份收入时所获得的幸福。同样，即便我们能够计算出15世纪所生产的绒面呢不断变动的价格，我们却始终不能知道它们的手感究竟如何。与其他社会科学家相比，事实上历史学家在计量方面特别受到限制；他不能询问他的研究对象有关其幸福或态度的问题，因此，也不能指望象心理学家和社会学家所能办到的那样，对这种幸福或政治态度作出哪怕是相对的计量。

尽管如此，人类过去经历的某些领域是无法计量的这一事实，决不能成为不去计量那些我们能够计量的领域的理由。至少，可以计量的领域有助于我们解释那些不可计量的领域。小阿瑟·施

莱辛格尔说：“几乎所有重要的问题之所以重要，正是由于它们不可能用计量结果来回答。”^①这样的观点忽视了下述事实，即如果不知道定量答案，我们也许就不能解释有关“重要”问题的证据。假如我们弄清了某人收入不断增加这一情况，那么，较之我们对其收入一无所知，就更容易解释此人不断增长的幸福。因此，即使我们原则上对“定性”问题要比对“定量”问题更感兴趣，这两个问题也是不可分割地联系在一起的。定量问题补充定性问题，定量证据也补充定性证据；其中任何一方不能取代另一方，任何一方也不能自称可以包罗历史研究的总体。不论其兴趣如何，一个历史学家的主要难题在于，他总是要面对一些不够充分的证据；我们始终不能完全肯定，我们的解释或描述是正确的。假如由于贬低计量法的重要性，历史学家完全排斥一切对定量证据的考虑，或把这种考虑降低到一种次要的位置，那么他就进一步削弱了他所获得的原已不充分的证据的意义。定量证据几乎肯定不能提供一个完整的答案，但它可以很好地提供答案的某些部分，而把它束之高阁，置之不理则是一种既不经济又不负责任的行为。

对计量史学还有一种更严厉，但同样是错误的批评，它的论点是使用计量法会导致过份简单化，会丧失有关过去的某些信息，会硬把个人纳入各种类别，结果使历史非人性化。当然，不管怎样使用分类或综合的方法，都会使人类历史经历的多样性简单化，因为那正是使用这些方法的目的。没有一个历史学家能够弄清楚这种多样性。历史学家的思维正象任何人的思维一样，在面对各种现象的复杂的多样性时，就不可避免地寻求某些类型和类似点，并舍弃或忘掉许多不符合这些类型的东西。与定性的印象主义史学相

① 小阿瑟·施莱辛格尔(A. Schlesinger jun.)：《人文主义者对经验主义的社会研究的考察》(The humanist looks at empirical social research)，载《美国社会学评论》(American Sociological Review)，第26卷(1961年12月)，第770页。

比，计量史学的优点在于，它的分类体系和方法、它所使用的假设和它所设定的类型，都是明确的，清晰的。要理解计量历史著作，人们无需深入了解历史学家的思路，也无需追寻他的思想进程，因为资料的分类和取舍都是展示在眼前的。在明确地探索类型和类似点时，计量史学家不得不总是承认他在进行简化，而且还必须描述他是怎样做的，他并不是无意识地清除各项不适合的证据。因此，计量史学家永远不会忽视历史证据固有的多样性，他所作的各种计量旨在把这种多样性变成一种可以理解的形式，同时也为那些偏离了所用的类型和平均值的证据指引一种研究方向。

确有一些质量不高的计量史学著作，其中的证据被硬纳入预先规定好的类别中，而且所使用的假设也与历史事实相违。然而，也不能说就没有质量很差的定性史学著作。可以肯定的是，尽管不是全部，但历史学家所作的陈述中确有许多是定量陈述，许多历史证据也是定量证据，得用计量技术来加以分析，而且使用计量方法的历史学家应该很好地运用这些技术。正象诠释中世纪的手稿或解释启蒙政治思想需要经验、训练和技能一样，处理数量方面的资料也要求具有一些特殊的分析方法和技术方面的知识。历史学家不能简单地研究一张数字表格，就希望立即理解它的意义，他必须学会从中抽引出其含义的技术，并把这种含义与他所收集到的另一些证据联系起来。因此，本书就是旨在帮助计量史学家很好地使用他的资料，并帮助那些阅读他的著作的人们去判断他是否这样做了。

下列各章是运用于历史问题和历史证据的一些计量方法。第一章涉及历史证据的分类和排比方法，以便使这种证据能被本书后面所叙述的方法加以分析。第二章讨论了一些以简明方式描述定量证据的方法，第三章则叙述了在分析证据时颇为有用的某些简单的数学方法（除了简单的中学算术和代数，并不要求具有更

高的数学知识)。因此,这三章就概括了历史学家在开始他的分析以前所必须进行的基本步骤。

第四章和第五章叙述了分析的初步阶段。在第四章中,讨论了运用图线和表格形式提供证据的一些方法,而在第五章中,这些提供证据的方法又通过讨论集中趋势(平均值)和偏差的计量而得到补充。在第六章中,所有这些技术都被运用于按编年(时间序列)排列的证据,并讨论了在时间序列的分析中特别重要的一些方法。

在第七章中,本书前几章研究的方法和概念被用来讨论统计方法,用以确定两个证据集合之间的关系。在此,相关关系的概念得到了讨论,并简单地介绍了最有力的统计技术之一——简单线性回归。

第八章涉及有关历史资料的一个特殊问题,即在传统的统计学教科书中常被忽视的证据短缺问题。在前几章讨论的方法的基础上,又提出了某些解决证据短缺问题的程序,并介绍了抽样概念。

最后,在第九章描述了某些计量分析的工具:电子计算机和电子计算器。本书前面各章讨论的方法并不要求使用这样的设备,不过事实表明,有了这些工具的帮助,计量分析可以更容易进行,时间也可以有所节省。

本书不是一本统计学的教科书,也不是要促进史学方法的讨论。它也不企图为计量历史资料分析中可能出现的所有问题提供答案。近年来,历史问题的计量研究已经获得了迅速的发展。这些研究使用了来源于其它各门社会科学的各种各样的方法,它们为数太多,在某些情况下也太复杂,所以不可能在此一一加以叙述。然而,这些研究和方法都具有一个共同的属于基本统计技术的核心,而这个核心部分将在后列各章中加以阐释。一个历史学家不论是想读一些运用统计分析所写的书和文章,或是他自己为

了写政治史、社会史、人口统计史、经济史甚至思想史时想运用定量性证据,都需要获得一定的技能,本书就是这些技能的一个导论。同时,它也是计量史学家最终可能要去阅读或参考的许多有关统计、计算、计量经济学或数学书籍的一个导论。

第一章 历史数据的分类

系统研究历史证据的首要要求之一,就是材料必须加以分类。历史学家自然要根据他所预定的概念和他所研究的对象,通过多种途径把材料加以分类。例如,历史学家们通常要把他们的材料分为第一手证据和第二手证据。第一手证据是产生于被研究的那个时期的证据,而第二手证据通常是另一些历史学家通过某种途径加工过的证据。历史学家用过的其它分类方法包括把证据区分为文字的和考古的,书写的和印刷的,或定量的和定性的。还可以用一些更为详尽的分类方法;例如,根据其来源把第一手证据加以分类,即把日记、法律诉讼状、法庭案卷、报纸、选举结果、商业档案都一一加以区分。

根据使用各种资料的经验,历史学家逐步制定出一些规则,使他们能够判断不同类型的资料的价值,并有效地加以利用。因此,他们把材料加以分类,部分原因是为了使这些规则便于运用。例如,巴格利把从1660年至1760年的教区记录分成两组:人头税统计册和教区对受洗、婚事、丧事的记录。把这两组加以区分后,巴格利认为,人头税统计册不是研究人口变动的可靠依据,而教区记录有时倒可能是很好的依据。^①

凡使用计量性材料的历史学家都必须不仅要学会根据材料来源及其可靠程度把材料加以分类,而且要学会按照如何表明这些

① 巴格利(J. J. Bagley):《历史的解释,2:从1540年至今日的英国历史资料》(Historical Interpretation, 2: Sources of English History, 1540 to the Present Day),英国哈蒙兹沃思1971年版,第84—154页。

材料在多大程度上适合于不同的分析方法来加以分类。他必须采取的的第一个步骤是仔细审查他的数据，而数据就是我们对他正在分析的材料称呼；然后通过一种有助于进行分析的方式把它们加以分类。我们所能运用的分类方式可分为列名分类、顺序分类和差别分类这样三种。我们能否按照其中的某一种方式把数据加以分类，取决于我们所掌握的信息或证据的多少。

(1) 列名分类

这第一种也是最简单的分类方式是用于通常的叙述的，在那种情况下我们通过给研究对象命名来把它们分成类别。通常，这是计算每一类别中包含多少研究对象的第一步。例如，〈英格兰土地勘查记录〉^①的编纂者在写到1086年位于肯特郡的瓦伊庄园时就是以列名数据的形式进行描述的：

“那里有52犁^②土地。在领主自营地上有9犁耕地；114户维兰^③和22户边地农占有17犁地。还有1座教堂，7名农奴；价值23先令8便士的4座磨坊；113英亩牧场和林地，可从那些土地的牧猪税中获得300头猪”。^④

在这一例子中，〈英格兰土地勘察记录〉的编纂者已考察了瓦伊庄园中的各种实物、人、牲畜和农用设施，给它们定了名称，并且

① 〈英格兰土地勘查记录〉(Domesday Book) 亦称“土地调查清册”，是英王威廉一世于1086年下令进行调查的全国土地、财产、牲畜和农耕人数的调查清册。——译者

② 原文为 Plough，按英史系地积单位。“犁”之大小因时期与地质不同而异，每“犁”从80英亩到100英亩。另一说约等于八头牛一年所耕之地积。——译者

③ 中世纪西欧封建社会的一种依附农民，在法国维兰有一定人身自由。——译者

④ 巴格利(J. J. Bagley)：《历史的解释》，1；英国中世纪史资料，1066—1540〉(Historical Interpretation, 1; Sources of English Medieval History, 1066—1540)，英国哈蒙滋沃思1965年版，第27页。