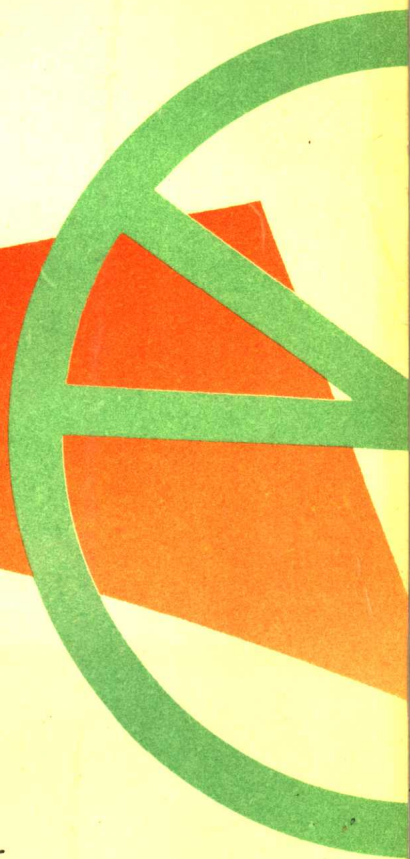
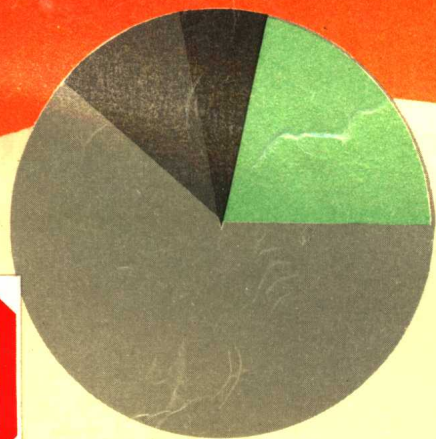


许飞琼
曾玉平

编著

统计学



中国统计出版社

C8
48

统 计 学

许飞琼 编著
曾玉平

中国统计出版社

(京)新登字 041 号

图书在版编目(CIP)数据

统计学/许飞琼,曾玉平编著. -北京:中国统计出版社,1995.9

ISBN 7-5037-1937-0

I. 统…

II. ①许… ②曾…

III. 统计学-概论

IV. C8

中国统计出版社出版

(北京三里河月坛南街 38 号 100826)

新华书店经销

北京市通县永乐印刷厂印刷

*

850×1168 毫米 32 开本 14 印张 35 万字

1995 年 9 月第 1 版 1995 年 9 月北京第 1 次印刷

印数:1—3500 册

※

ISBN 7-5037-1937-0/C·1128

定价:19.00 元

(版权所有 不得翻印)

前 言

在人类社会的发展史上,统计是伴随着生产活动和国家管理的需要而产生和发展起来并逐步成为一门科学的。进入当代社会后,统计作为人们认识社会的有力武器,更是在社会经济、社会政治和社会生活乃至科学研究等一切领域中得到了非常广泛的应用,起着日益重要的作用。大至国家的宏观决策,小至企事业单位的微观管理和城乡人民的家居生活,都离不开对统计的应用。因此,研究统计的科学即统计学,便构成一门通用方法论科学。

本书是一部统计学原理著作,在介绍一般统计学知识的同时,亦揉进了作者学习、研究统计学的心得。因此,它具有以下几个特点:一是具有全面性和系统性,不仅介绍了传统统计学中最基本的理论与方法,而且补充了一些最新的、在实际工作中运用很广泛的统计理论与方法;二是力图用较通俗的语言来阐述较深奥的统计概念,用表格、图形等形式直观而形象地描述统计学的基础理论;三是注重用现实生活中的各种实例来研讨、分析统计问题。这样做的目的,在于使读者学习统计学知识、研究统计学问题的兴趣得到提高。

本书的写作过程历经三年,作者曾进行过多方面的统计调查研究,因此,本书可以说是理论与实践相结合的产物。全书分十五章,其中第一、二、三、四、五、六、七、八、九、十一章由武汉大学许飞琼撰写,第十、十二、十三、十四、十五章由国家统计局曾玉平撰写,全书由许飞琼统稿、定稿。

本书主要适用于高校师生和统计工作者。在本书脱稿并付梓之际,作者希望得到同行专家、学者和读者的批评指正。

作者 谨启

一九九五·四

目 录

第一章 导论	(1)
第一节 统计学概述	(1)
第二节 统计学的产生与发展	(6)
第三节 统计学中的术语	(15)
第四节 统计学的基本方法	(20)
第五节 统计体制与统计活动	(23)
第二章 统计资料搜集方法	(26)
第一节 统计资料及其搜集方式	(26)
第二节 统计调查方案的制订	(32)
第三节 统计误差	(37)
第三章 统计资料的整理	(39)
第一节 统计资料整理概述	(39)
第二节 频数分布	(45)
第三节 统计表和统计图	(58)
第四章 统计集中趋势与离散趋势	(70)
第一节 集中趋势及其指标计算	(70)
第二节 离散趋势及其指标计算	(93)
第三节 偏度和峰度	(104)
第五章 时间序列分析	(111)
第一节 时间序列分析概述	(111)
第二节 时间序列分析指标	(115)
第三节 时间序列的变动因素	(124)
第四节 长期趋势的测定	(128)
第五节 季节变动的测定	(138)
第六节 循环变动的测定	(148)

第六章	指数	(151)
第一节	指数的意义和种类	(151)
第二节	指数的编制方法	(154)
第三节	指数序列及指数体系分析	(168)
第七章	概率分布	(178)
第一节	概率及其计算法则	(178)
第二节	随机变量及概率分布	(188)
第三节	二项分布和普阿松分布	(193)
第四节	正态分布	(198)
第八章	抽样法	(205)
第一节	概述	(205)
第二节	抽样技术	(210)
第三节	抽样误差	(216)
第四节	抽样推断	(231)
第九章	假设检验	(243)
第一节	假设检验概述	(243)
第二节	总体平均数的假设检验实例	(251)
第三节	总体成数的假设检验实例	(256)
第十章	非参数分析法	(260)
第一节	正负符号检验法	(260)
第二节	威尔科克森符号秩检验法	(265)
第三节	游程检验法	(271)
第四节	χ^2 检验法	(273)
第五节	中位数检验法	(277)
第十一章	回归分析与相关分析	(280)
第一节	概述	(280)
第二节	一元回归分析	(285)
第三节	相关分析	(292)
第四节	曲线回归分析	(301)
第五节	多元回归分析	(307)
第六节	自回归分析	(313)

第七节 回归分析与相关分析应注意的问题	(316)
第十二章 边际与弹性分析	(318)
第一节 边际分析	(318)
第二节 弹性分析	(326)
第十三章 聚类分析	(338)
第一节 聚类分析概述	(338)
第二节 数据变换处理	(339)
第三节 聚类统计量	(342)
第四节 聚类方法	(347)
第十四章 统计预测	(360)
第一节 统计预测概述	(360)
第二节 统计预测方法	(363)
第三节 统计模型预测	(376)
第十五章 统计决策	(382)
第一节 统计决策概述	(382)
第二节 非概率决策分析	(386)
第三节 先验概率决策分析	(390)
第四节 后验概率决策分析	(394)
第五节 决策树分析	(396)
附表 1 常用对数表	(402)
附表 2 反对数表	(407)
附表 3 二项分布累积概率表	(414)
附表 4 普阿松分布用的 $e^{-\lambda}$ 值表	(418)
附表 5 标准正态分布表	(420)
附表 6 t 分布表	(422)
附表 7 相关系数临界值表	(425)
附表 8 F 分布表	(428)
附表 9 威尔科克森 T 值表	(432)
附表 10 威尔科克森秩和检验临界值表	(434)
附表 11 游程检验表	(436)

附表 12 χ^2 分布表 (438)

附录 主要参考书目 (440)

第一章 导 论

第一节 统计学概述

一、统计学的涵义

统计学(英文为 Statistic)是以社会现象和自然现象的数据表现为研究对象,并需要结合社会科学和自然科学的一门通用方法论科学。一方面,随着科学实验的发展,定量分析在各个领域都得到了日益广泛的应用,数据成了检验理论的科学依据,于是,统计学的理论与方法就具有了普遍的适用性和方法论意义;另一方面,作为一门交叉学科,统计学又与其他学科相结合,使各门应用统计学得以产生,如经济统计学、人口统计学、环境统计学、工程统计学、科技统计学等。此外,统计学与其他学科相结合,还衍生出一些新的学科,如经济计量学就是统计学与经济学的相互结合而产生的一门新学科。因此,统计学在学科性质上应是一门交叉于社会科学和自然科学并具有综合性、横断性与边缘性的学科。

从现代统计学的研究角度出发,统计学有理论统计学与应用统计学之分。理论统计学主要是从原理角度探讨统计方法,证明各统计公式的来龙去脉,其特点是着重于统计数学原理的推导。理论统计学的统计理论一般可以分为 16 个研究大类:数学方法、概率论、概率分布、抽样分布、估计、假设检验、关联性、相依性、回归分析、方差分析、抽样、试验设计、随机过程理论、随机过程推断、运筹学、应用专

题、一般特点^①。而应用统计学则是以理论统计学为基础,结合某一领域或专业现象的特点,来研究、认识所研究现象的具体理论与方法。如社会经济统计学、管理统计学、科技统计学、物理统计学、环境统计学等,就均属于应用统计学。

从现代统计学的技术角度出发,统计学又分为描述统计学和推断统计学。所谓描述统计学,是指运用指标、图和表等对所搜集的资料加以整理和分析以研究问题的理论方法。如1994年我国国内生产总值达43800亿元,按可比价格计算,比1993年增长11.8%,就是用指标来描述我国国内生产总值的发展情况;再如,某股份有限公司要向股东表明公司历年来产品销售收入迅速增长的情况,而采用了图1-1的曲线图来描述。

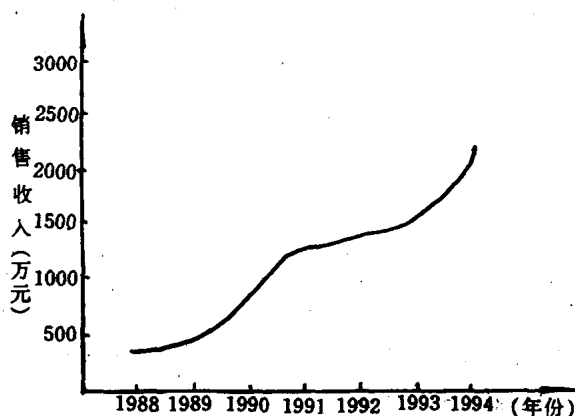


图 1-1

所谓推断统计学也称归纳统计学,是指通过样本来推断总体的方法。由于客观现象的复杂性,人们不可能对总体全部单位进行资料的搜集描述,即使有时可以对总体全部单位进行资料的搜集描述,但因工作量大,也通常不对全部单位的资料进行搜集描述,而只是从总

^① 国际统计学会:《统计理论与方法文摘》1991年第32卷2期。

体全部单位中抽取一部分单位进行观察,用对抽取部分观察的结果去推断总体,因此,推断统计在统计学中有着十分重要的意义。一般而言,推断统计必须以描述统计所用的方法分析样本资料,并且是根据描述样本的结果再进一步从局部推断全体,从已知推断未来。当然,样本作为总体的一部分,包含的信息不完备,其结果常含有某种成分的不确定性(这种不确定性因样本推断总体所产生的误差而产生),但是推断统计根据概率论的原理可以使归纳推断所产生的不确定性得到度量。

作为一级学科,统计学是一个独特而庞大的学科体系,其十分丰富的内容可以用图 1-2 来表示。

二、统计学界定之争

在我国统计学理论的发展中,不同的学者对统计学的界定亦存在着很大差异。具体而言,可以概括为以下四种观点:

1. 数理统计学论。该种观点认为,统计学就是数理统计学,数理统计学可以广泛应用于自然技术界和经济领域,它是一门通用的方法科学。它以自然现象和社会现象的数量方面为研究对象。在西方国家,自 20 世纪 40 年代以后,社会统计学及社会统计学派逐渐被数理统计学及其学派所取代,因而一提到统计学,就都理解为数理统计学。至于工业统计学、商业统计学、管理统计学等学科,都是数理统计方法在各领域中的应用。

2. 社会经济统计学论。该种观点认为,统计学是一门独立的社会科学,数理统计学是数学的一个分支,不能属于统计学。社会经济统计学以历史唯物主义和马克思主义政治经济学为理论基础。

3. 并存发展论。该种观点认为,数理统计学与社会经济统计学是并存的 two 门不同的独立的统计学,它们各有不同的研究对象、内容和方法,两者都产生于实践,并对实践起着一定的指导作用。

4. 大一统论。该种观点认为,统计学是一门统一的通用的方法论科学,数理统计学和社会经济统计学是它的两个分支,即不过是统计

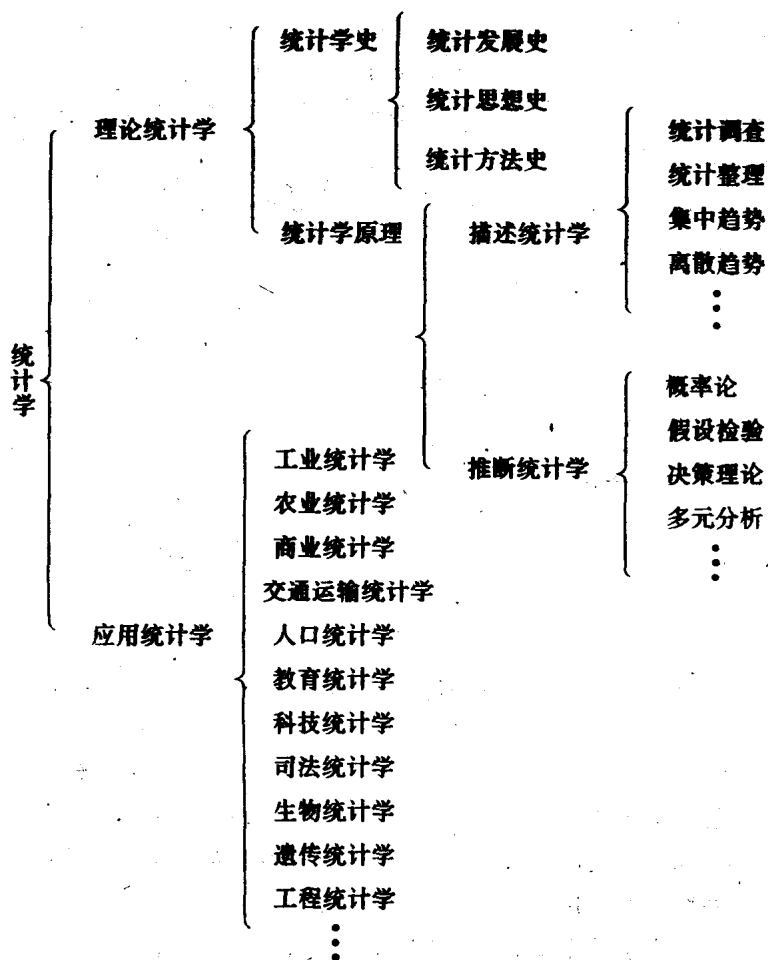


图1- 2 统计学学科体系结构

学理论与方法在自然领域和社会领域的应用而已。

在上述观点中,主张社会经济统计学是独立的学科论者对该学科又存在着不同的看法:一种观点认为社会经济统计学是一门实质性科学,是以大量社会经济现象的数量方面为研究对象的社会科学,目的是找到社会经济发展规律在具体地点、时间、条件下的数量表现;第二种观点认为社会经济统计学是方法论科学,而不是实质性科

学,它是对社会经济现象数量方面进行调查研究的方法论科学;第三种观点认为社会经济统计学是方法论与实质性兼而有之的科学,理由是社会经济具有社会与经济专业的特征,必然要涉及专业的特有方法与实质性科学的相关知识,故社会经济统计学是实质与方法的统一体。

三、统计学的特点

统计学之所以成为一门独立的学科,在于它具有与其他学科不同的特点。其特点主要体现在以下几个方面:

1. 数量性。这是统计学的首要特点。统计学研究的是现象的数量方面,它通过对各种现象数量的收集、整理和分析来认识现象。如数量的多少、现象之间的数量关系、质量互变的数量界限,等等。

2. 总体性。统计学研究的目的不仅仅是反映个别现象或现象的个别数量表现,而是要通过数据来描述和揭示由大量单个现象所构成的总体的数量规律性。例如,要分析某地区职工的收入水平,由于每个职工的收入水平有差异,若用某职工的收入水平来衡量该地区的收入水平显然是不科学的,只有通过对该地区所有职工收入水平的分布状态进行描述,以及计算得出其平均水平才能得到规律性的东西。再如,人口死亡率、种子发芽率、空气污染程度,等等,无一不以大量单个现象所构成的总体来进行研究。其实,规律性的结论都必然要通过对大量个别现象的观察、记录和分析才能获得。因此,统计学是对自然现象和社会现象总体的定量认识的科学。

3. 具体性。统计学研究的不是空洞的理论、不是抽象的数量研究(抽象的数量研究即是数学),而是采用具体的方法对具体地点、具体时间、具体条件下的现象的数量进行研究。例如,对1993—1994年度的零售物价水平进行研究得出通货膨胀率在20%以上的结论,即是统计学研究的具体性表现。

4. 广泛性。统计学的研究目的是为了应用,现代统计学的理论与方法不仅应用于社会领域,而且应用于自然领域。如应用于经济、政

治、法律、物理、生物、医学、农业等各个领域,从而使统计学具有广泛性的特征。

5. 工具性。统计学是方法论科学,它是通过自己的统计方法和统计方法论的研究而服务于数量关系的研究的。统计学以自己独有的方法来揭示现象数量变化的规律。例如,借助于统计学可以揭示出一批产品是否符合标准及不符合标准的状况,但它不去研究产品不合标准的原因,因为后者属管理学的内容。由此可知,统计学只是一门方法论科学,借助于统计学独特的方法对社会现象和自然现象进行反映,以探索现象发展变化的情况和规律。这样,统计学便成了达到某种研究目的的工具而具有工具性特征。

第二节 统计学的产生与发展

纵观历史,统计学的产生与发展大致可以划分为古典统计学、近代统计学、现代统计学三个阶段。

一、古典统计学

古典统计学,是指 17 世纪中叶到 18 世纪中叶这一百年间,处于萌芽状态的统计学。

“统计”一词来源已久,最初语源出于中世纪拉丁语的 status,意即各种现象的状态和状况。由这一语根组成的意大利语 stato,表示国家的概念,以及表示关于各国的国家结构和国情这一方面知识的总称。

在 13 世纪至 16 世纪中叶,由于资本主义发展的不平衡,各国国情的不一致,欧洲主要国家都深深感到有调查国内、国外情况的必要,于是,从意大利开始,各国相继进行本国和他国的历史沿革、地理条件、国家的典章制度、财政收入和军事力量、居民风俗习惯、国家的工商业、交通运输等国情调查。当各国调查资料积累到一定数量时,于是有人将之汇编成册,并从事初步研究。如德国学者芒斯特(Se-

bastian Münster, 1489—1552)所编的《世界志》,即为第一部反映这些国家财富调查的科学统计著作^①。此外,意大利学者桑苏汶诺(Francesco Sansovino, 1521—1586)所编的《国家制度志》、德国学者塞肯道夫(Veit Ludwig Von Seckendorff, 1626—1692)所编的《德意志王国志》、以及英国学者张伯伦(Edward Chamberlayne, 1618—1703)等编的有些书目,都是关于世界各国或本国的国情方面的著作。这些著作一方面(主要是德、法派)为当政者或准备从政者提供了必要的国内外知识;另一方面(主要是英、荷派)为国家制定方针政策以及资产阶级在国内外经商、争夺海外市场和开拓殖民地提供了一定的国内外知识。由于这些著作主要是论述和反映国情方面的知识,所以后人称之为“国势论”。这一时期“国势论”的研究方法特色,主要是简单的文字记述,而缺乏必要的比较、因果等分析方法。

进入17世纪中叶以后,国势论的研究对象、研究方法,尤其是数量计算方法,较之17世纪中叶以前得到了一定的发展。如德国学者康令(Hermann Conring, 1606—1681),除采用记述方法外,还采用了分组法、比较法,这在他的《国势学》著作中作了详细的反映。尽管康令在《国势学》中不是具体采用数据计算和比较,而仅仅采用笼统的比较说法(如甲国比乙国人口多,至于多多少则没有数据表述),但其研究方式得到当时许多学者的欢迎和后来学者的继承,并形成了一个独有的派别——国势学派,也称记述学派。如施梅泽尔(Martin Schmeitzel, 1679—1747)在耶拿、哈大两所大学讲授的《政治统计学讲义》,则明显继承了康令的学说,且开始使用“Statisticum”(统计学)一词^②。只是施梅泽尔仅仅继承了康令学说,对国势学说的发展并无实质上的贡献。有突出贡献的则是康令的学生阿亨瓦尔(Gottfried Achenwall, 1719—1772)。

阿亨瓦尔是国势学派的主要代表人物,他在其著作《近代欧洲各

^① 《统计学的历史·原理和方法》,1891年,德国统计学家迈著,高庆丰著,《欧美统计学史》,中国统计出版社,1987年版。

^② 韦斯特加德著:《统计学史》1932年,英文版,第7页。

国国势学纲要》中,首次正式用了“统计学”(Statistik)一词,并把它当作学名使用。他将统计学定义为“国家显著事项学”,是关于国家组织、人口、军队、领土、居民职业、地面地下资源的学问。认为统计学研究的目的一方面是为从政者提供必要的有关国家管理的知识;另一方面是用来比较探索国家盛衰的因果关系,作为统治者兴利除弊的借鉴。只是阿亨瓦尔的“统计学”论的研究方法仅仅是以文字记述,与康令一样没有借用数据资料去探索现象的规律性,从而仅有“统计学”之名,而无“统计学”之实。

有“统计学”之实的,则是与国势学派同期的政治算术学派。16世纪末到17世纪中,英国连续几次遭瘟疫,居民生活和卫生出现了严重危机,加之荷、法两国又趁机频频进攻,英国人对国家的现状和前途产生了无限的忧虑和悲观。为此,一些爱国人士为了救人民、救国家,开始大量搜集数据资料,并用图表、分组、对比等方法撰写文章和著作来证明英国能够战胜荷、法两国并走向富强。约翰·格朗特和威廉·配第及他们的著作则是最杰出的代表。

约翰·格朗特(John Graunt, 1620—1674), 1662年出版了一本题为《对死亡表的自然观察和政治观察》的书,该书以人口问题为中心,通过对客观现象中的数量关系(如男、女性别关系,男、女出生率等),揭示出一系列的统计规律,如,每一百人中死亡人数与各种年龄有一定,的统计规律性;一切疾病和事故在全部死亡原因中占有稳定的百分比,而传染病所占的百分比波动很大;等等。在这里,格朗特已经认识到大量观察法的重要性,因此,《对死亡表的自然观察与政治观察》一书后来被许多统计学家,尤其是英国数理统计学派的统计学家誉为“真正统计科学的肇端”。

威廉·配第(William Petty, 1623—1687), 1675年左右写了一本题为《政治算术》的著作。在该书中,他对统计学的研究目的、任务和对象作了初步的阐述。配第认为社会现象和自然现象一样存在着客观规律,而“政治算术”则可以作为探索规律和法则的手段。在书中,配第运用了数字、重量、尺度、图表以及分组方法、推算方法来研究

社会经济现象。只是配第论述问题所涉及的理论与方法，他都冠之以“政治算术”，而不是“统计学”，因此，他的这种政治算术学观点有“统计学”之实，但无“统计学”之名。然而，威廉·配第的理论与方法为近代统计学奠定了一定的基础。马克思曾在《反杜林论》中写道：“配第创造的‘政治算术’即一般所说的统计”^①。在对配第的评价方面，马克思在《资本论》第一卷中更明确的写道：“配第是政治经济学之父，在某种程度上也可以说是统计学的创始人”^②。

以“统计学”之名取代“政治算术”，则是18世纪中末叶到19世纪初这个阶段。一方面，英国将德语 *statistik* 翻译成了 *statistic* (统计学)；另一方面，德国社会统计学家克尼斯(K·G·A·Knies)在1850年提出把统计学的名称转让给政治算术派。从此以后，才真正使统计学成为有政治算术内容的统计学。

格朗特与配第的学说形成了“政治算术学派”。随着资本主义的日趋发展，政治算术学派很快在欧洲各国传播，并在17世纪末叶以后，由各自研究对象的重点分出两大派别，即人口统计派和经济统计派。人口统计派的奠基人是英国天文学家哈利(Edmund Halley, 1656—1742)，其《根据勃瑞斯劳城出生、死亡统计表作出的人类死亡率估计》一文，提出了著名的“哈利生命表”^③。经济统计派的代表人物是英国经济学家兼统计学家金氏(Gregory King, 1648—1712)和法国著名元帅沃班(S·Vanban, 1633—1707)，前者代表作为《国家收入论》，后者的代表作为《王国什一税论》。到18世纪，政治算术则得到了更进一步的发展，贡献最突出的是德国的苏斯米尔希(Johann Peter Süssmilch, 1707—1767)，他对政治算术作了非常系统的论述，在其代表作《神定秩序》一书中，提出了大量观察法，证明了大数定律的存在，这在统计学的发展上具有科学的价值。

① 《苏联1965年国民经济·统计年鉴》第10页，莫斯科，中国统计出版社。

② 《列宁全集》，第3卷第516—517页，人民出版社，1959年版。

③ V·约翰著、足利末男译《统计学史》第207页，1956年日文版，K·皮尔逊《十七、十八世纪统计学史》第79页，1978年英文版。