



21st CENTURY
实用规划教材

21世纪全国高职高专财经管理系列实用规划教材

统计学原理

主 编 廖江平 刘登辉



北京大学出版社
PEKING UNIVERSITY PRESS



中国农业大学出版社
CHINA AGRICULTURAL UNIVERSITY PRESS

21 世纪全国高职高专财经管理系列实用规划教材

统计学原理

主 编 廖江平 刘登辉



北京大学出版社
PEKING UNIVERSITY PRESS



中国农业大学出版社
CHINA AGRICULTURAL UNIVERSITY PRESS

内 容 简 介

本书根据高职高专的教学特点,参照教育部最新制定的《高职高专教育基础统计课程教学基本要求》,借鉴和吸收有关统计学研究的最新成果,并在作者总结多年统计学教学经验的基础上编写而成。

本书共分10章,主要包括4方面的内容:一是统计调查及统计整理的意义和方法;二是主要统计指标及指数的作用和计算方法;三是时间数列、抽样推断及相关与回归分析的特点和方法;四是国民经济核算的相关内容及指标。本书概念准确,层次清晰,重点突出,注重定量分析与定性分析的有机结合。

本书可作为高职高专院校、本科院校的二级学院及民办高校的财经管理和经济贸易类专业的教材,也可作为从事实际经济工作人士的学习参考资料。

图书在版编目(CIP)数据

统计学原理/廖江平,刘登辉主编. —北京:中国农业大学出版社;北京大学出版社,2009.8

(21世纪全国高职高专财经管理系列实用规划教材)

ISBN 978-7-81117-825-8

I. 统… II. ①廖…②刘… III. 统计学—高等学校:技术学校—教材 IV. C8

中国版本图书馆CIP数据核字(2009)第124701号

书 名: 统计学原理

著作责任者: 廖江平 刘登辉 主编

总 策 划: 第六事业部

执行策划: 吴 迪

责任编辑: 卢 东 魏秀云

标准书号: ISBN 978-7-81117-825-8

出 版 者: 北京大学出版社(地址: 北京市海淀区成府路205号 邮编: 100871)

网址: <http://www.pup.cn> <http://www.pup6.com> E-mail: pup_6@163.com

电话: 邮购部 62752015 发行部 62750672 编辑部 62750667 出版部 62754962

中国农业大学出版社(地址: 北京市海淀区圆明园西路2号 邮编: 100193)

网址: <http://www.cau.edu.cn/caup> E-mail: cbsszs@cau.edu.cn

电话: 编辑部 62732617 营销中心 62731190 读者服务部 62732336

印 刷 者: 北京飞达印刷有限责任公司

发 行 者: 北京大学出版社 中国农业大学出版社

经 销 者: 新华书店

787毫米×980毫米 16开本 16.75印张 328千字

2009年8月第1版 2009年8月第1次印刷

定 价: 25.00元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话: 010-62752024

电子邮箱: fd@pup.pku.edu.cn

21 世纪全国高职高专财经管理系列实用规划教材(已出版)

序号	标准书号	书 名	主 编	出版时间
1	ISBN 978-7-5038-4836-0	统计学原理与实务	姜长文、简家进	2007 年 8 月
2	ISBN 978-7-5038-4846-9	经济法原理与实务	孙晓平、邓敬才	2007 年 8 月
3	ISBN 978-7-5038-4835-3	现代公共关系原理与实务	张美清	2007 年 8 月
4	ISBN 978-7-5038-4847-6	广告原理与实务	郑小兰、谢 璐	2007 年 8 月
5	ISBN 978-7-5038-4841-4	管理学原理	季 辉、冯开红	2007 年 8 月
6	ISBN 978-7-5038-4858-2	企业管理	张 亚、周巧英	2007 年 8 月
7	ISBN 978-7-5038-4851-3	人力资源管理	李蔚田、傅 航	2007 年 11 月
8	ISBN 978-7-5038-4859-9	市场营销学	李世宗、李建峰	2007 年 8 月
9	ISBN 978-7-5038-4845-2	基础会计教程与实训	李 洁、王美玲	2007 年 8 月
10	ISBN 978-7-5038-4849-0	基础会计原理与实务	侯旭华、缙宇英	2007 年 8 月
11	ISBN 978-7-5038-4843-8	审计学原理与实务	马西牛、杨印山	2007 年 8 月
12	ISBN 978-7-5038-4860-5	财务会计	李 哲、孙晓平	2007 年 8 月
13	ISBN 978-7-5038-4848-3	税务会计实用教程	李克桥、郭 华	2007 年 8 月
14	ISBN 978-7-5038-4837-7	财务管理教程与实训	张 红、景云霞	2007 年 8 月
15	ISBN 978-7-5038-4853-7	会计电算化实用教程	张耀武、卢云峰	2007 年 11 月
16	ISBN 978-7-5038-4839-1	国际金融基础与实务	冷丽莲、刘金波	2007 年 8 月
17	ISBN 978-7-5038-4852-0	国际贸易理论与实务	程敏然、贺亚茹	2007 年 8 月
18	ISBN 978-7-5038-4838-4	货币银行学	曹 艺、卞桂英	2007 年 8 月
19	ISBN 978-7-5038-4842-1	进出口贸易实务	周学明、金 敏	2007 年 8 月
20	ISBN 978-7-5038-4840-7	财政基础与实务	才凤玲、张云莺	2007 年 8 月
21	ISBN 978-7-5038-4856-8	财政与金融	谢利人、郝巧亮	2007 年 8 月
22	ISBN 978-7-5038-4844-5	国际结算	徐新伟	2007 年 8 月
23	ISBN 978-7-5038-4855-1	商品学概论	方凤玲、杨 丽	2007 年 8 月
24	ISBN 978-7-5038-4850-6	商务谈判	范银萍、刘 青	2007 年 8 月
25	ISBN 978-7-5038-4854-4	现代物流管理	沈 默、李承霖	2007 年 8 月
26	ISBN 978-7-5038-4857-5	仓储与配送管理实务	郭曙光	2008 年 2 月
27	ISBN 978-7-301-11632-6	电子商务实务	胡华江、余诗建	2007 年 2 月
28	ISBN 978-7-301-09440-6	会计电算化实用教程(第 2 版)	刘东辉	2008 年 5 月
29	ISBN 978-7-5038-5012-7	会计英语	杨 洪	2008 年 2 月
30	ISBN 978-7-5038-5017-2	基础会计实训教程	王桂梅	2008 年 3 月

序号	标准书号	书 名	主 编	出版时间
31	ISBN 978-7-5038-5027-1	财务会计实用教程	丁增稳、高 丛	2008 年 5 月
32	ISBN 978-7-81117-529-5	中小企业管理	吕宏程	2008 年 8 月
33	ISBN 978-7-81117-532-5	国际商务谈判	卞桂英、刘金波	2008 年 8 月
34	ISBN 978-7-81117-537-0	金融英语	刘 娣	2008 年 8 月
35	ISBN 978-7-81117-542-4	成本会计实训教程	贺英莲	2008 年 8 月
36	ISBN 978-7-81117-541-7	营销策划技术	方志坚	2008 年 8 月
37	ISBN 978-7-81117-570-7	市场调研案例教程	周宏敏	2008 年 9 月
38	ISBN 978-7-81117-590-5	秘书理论与实务	赵志强	2008 年 11 月
39	ISBN 978-7-81117-592-9	成本会计	李桂梅	2009 年 2 月
40	ISBN 978-7-81117-645-2	资产评估	董亚红	2009 年 2 月
41	ISBN 978-7-81117-661-2	消费心理学	臧良运	2009 年 2 月
42	ISBN 978-7-81117-759-6	零售学	陈文汉	2009 年 7 月
43	ISBN 978-7-81117-802-9	管理信息系统	刘 宇	2009 年 7 月
44	ISBN 978-7-81117-756-5	统计学基础	阮红伟	2009 年 8 月
45	ISBN 978-7-81117-753-4	基础会计教程	侯 颖	2009 年 8 月
46	ISBN 978-7-81117-806-7	现代企业管理	于翠华、贾志林	2009 年 8 月
47	ISBN 978-7-81117-825-8	统计学原理	廖江平、刘登辉	2009 年 8 月
48	ISBN 978-7-81117-675-9	经济法实用教程	胡卫东、吕 玮	2009 年 8 月

电子书(PDF 版)、电子课件和相关教学资源下载地址: <http://www.pup6.com/ebook.htm>, 欢迎下载。

欢迎免费索取样书, 请填写并通过 E-mail 提交教师调查表, 下载地址:

<http://www.pup6.com/down/教师信息调查表 excel 版.xls>, 欢迎订购。

联系方式: 010-62750667, wudi1979@163.com, linzhangbo@126.com, 欢迎来电来信。

21 世纪全国高职高专财经管理系列实用规划教材

专家编审委员会

主 任 刘诗白

副 主 任 (按拼音顺序排名)

曹志平 方凤玲 高建宁 侯旭华 吉文丽

季 辉 金绍珍 李世宗 刘晓玉 张美清

委 员 (按拼音顺序排名)

程敏然 陈兆芳 丁增稳 范银屏 付丽红

何惠珍 计桂贤 靳 生 李华翔 李 洁

李克桥 李 哲 刘春朝 刘建民 吕 玲

彭 杰 沈 萍 孙希月 万久玲 王庆春

王 潇 宣国萍 杨丽娜 杨印山 姚 伟

姚 旭 周仁贵 田朝阳 银加峰 黄 浩

法律顾问 杨士富

丛书总序

随着我国改革开放的逐步深化,社会主义市场经济对高等职业技能型人才的需求迅猛增加。2002年,《国务院关于大力推进职业教育改革与发展的决定》(国发[2002]16号)的颁布,揭开了我国高等职业教育发展的新篇章。为贯彻落实《国务院关于大力推进职业教育改革与发展的决定》,“十一五”期间,教育部、财政部决定实施国家示范性高等职业院校建设计划,通过重点建设100所国家示范性高职院校,带动全国高职院校深化改革,提升高等职业教育的整体水平。国家启动示范性高等职业院校建设计划,标志着我国高等教育进入了一个追求内涵发展的新的历史阶段,这是科学发展观在我国高等教育领域的具体体现,对促进我国高等职业教育更好更快地发展具有巨大的战略意义。

财经管理类专业是我国高职高专教育极其重要的组成部分。2005年,全国高职高专院校在校生427万,其中财经管理类专业在校生超过80万,占18.7%。高职高专财经管理类专业主要着眼于培养社会主义市场经济发展所需要的德、智、体全面发展的高素质专门人才,要求学生具有较强的职业技能和较好的创新精神以及实践能力。

在当前开拓新型工业化道路,推进全面小康社会建设的新时期,进一步加强经济管理人才的培养,注重经济理论的系统化学习,特别是现代经济管理理论的学习,提高学生的专业理论素质和应用实践能力,培养出一大批高水平、高素质的经济管理人才,越来越成为提升我国经济竞争力、保证国民经济持续健康发展的重要前提。这就要求高职高专财经管理类专业职业教育要更加注重依据国内外社会经济条件的变化,适时变革和调整教育目标和教学内容;要求财经管理学科专业更加注重应用、注重实践、注重规范、注重国际交流;要求财经管理学科专业与其他学科专业相互交融与协调发展;要求财经管理职业教育培养的人才具有更加丰富的社会知识和较强的人文素质及创新精神。要完成上述任务,高职高专院校需要进行深入的教学改革和创新,特别是要搞好有较高质量的教材的编写和创新工作。

出版社的领导和编辑通过对国内高职高专院校财经管理学科教材使用情况的调研,在与各院校的专家学者讨论的基础上,决定组织编写和出版《21世纪全国高职高专财经管理系列实用规划教材》,这是一项有利于促进高职高专院校财经管理专业教学改革发展的重要措施。

本系列教材是按照高职高专院校财经管理学科专业规范、培养方案,以及课程教学大纲的要求,合理定位,由长期在教学第一线从事教学工作的教师立足于21世纪财经管理学

科发展的需要,深入分析财经管理专业学生现状及存在问题,探索财经管理专业学生综合素质培养的途径,以科学性、先进性、系统性和实用性为目标,其编写的特色主要体现在以下几个方面。

(1) 关注财经管理学科发展的大背景,在掌握必要的理论知识基础上,着眼于增强教学内容的联系实际和应用性,突出创造能力和创新意识。

(2) 体系完整、严密。本系列涵盖财经类、管理类相关专业,并把握相关课程之间的关系,整个系列丛书形成一套完整、严密的知识结构体系。

(3) 内容新颖。借鉴国内外最新的教材,融会当前有关财经管理学科的最新理论和实践经验,用最新知识充实教材内容。

(4) 合作交流的成果。本系列教材是由全国上百所高职高专院校教师共同编写而成,在相互进行学术交流、经验借鉴、取长补短、集思广益的基础上,形成编写大纲。最终融合了各地特点,具有较强的适应性。

(5) 案例教学。教材具备大量案例研究分析,让学生在学习过程中理论联系实际,特别列举了我国财经管理工作中的大量实际案例,这可大大增强学生的实际操作能力。

(6) 注重能力培养。力求做到不断强化自我学习能力、思维能力、创造性解决问题的能力以及不断自我更新知识的能力,促进学生向着富有鲜明个性的方向发展。

(7) 配套产品种类丰富。每本教材除了有电子课件方便老师备课以外,还提供有教材习题答案、考试题库,为使用本系列教材的老师提供了方便。

作为高要求,高职高专财经管理类教材应在基本理论上做到以马克思主义为指导,结合我国财经管理工作的新实践,充分汲取中华民族优秀文化和西方科学管理思想,形成具有中国作风、气派和特色的创新教材。这一目标不可能一蹴而就,需要作者通过长期艰苦的学术劳动和不断地进行教材内容的更新才能达成。我们希望这一系列教材的编写,将是我国拥有较高质量的高职高专财经管理学科教材建设工程的新尝试和新起点。

我们要感谢参加本系列教材编写和审稿的各位老师所付出的大量卓有成效的辛勤劳动。由于编写时间紧、相互协调难度大等原因,本系列教材肯定还存在一些疏漏。我们相信,在各位老师的关心和帮助下,本系列教材一定能不断地改进和完善,并在我国高职高专财经管理类学科专业的教学改革和课程体系建设中起到应有的促进作用。

《21 世纪全国高职高专财经管理系列实用规划教材》

专家编审委员会

2007 年 8 月

前 言

进入 21 世纪,知识的更新速度不断加快,社会对知识的需求也日益增长。无论是企业经营管理与决策还是科学研究,都越来越依赖统计分析,统计分析方法已成为各领域进行分析与研究的基本方法。因此,统计分析不仅是人们认识社会经济现象的有力武器,而且还是政府管理经济的有力工具,更是企业经营者进行科学管理与决策的重要方法。

本书是编者在总结多年的统计学教学、实践经验,并吸收有关统计学研究的最新成果和国内外同类优秀教材的基础上,参照教育部最新制定的《高职高专教育基础统计课程教学基本要求》编写而成。本书在编写过程中既注重基本理论的介绍,又注重统计方法的实际应用,力求概念准确、层次清晰、重点突出、简明扼要、通俗易懂;同时,针对高职高专教育的特点,注重理论知识与实际应用的统一。考虑到近年来统计理论与实践的发展趋势及应用热点,本书适当引入了一些新方法和新概念,并以阅读材料、统计小辞典的形式补充了较多的案例和相关知识以扩展读者的知识面,提高读者的学习兴趣。每章后附有适量课后习题,供学生巩固所学知识,提高自学能力之用。

本书由廖江平、刘登辉主编。第 1 章、第 2 章、第 3 章、第 4 章、第 5 章由江苏经贸职业技术学院的廖江平编写;第 6 章、第 7 章、第 8 章、第 9 章、第 10 章由江苏经贸职业技术学院的刘登辉编写。本书在编写过程中参阅了国内外一些重要文献,引用了大量最新经济统计数据,并得到了何滔滔高级会计师、徐立伯博士的大力支持,在此一并表示衷心的感谢。

由于编者学识水平有限,加之时间仓促,书中难免有疏漏之处,恳请广大读者批评指正。

编 者

2009 年 4 月

目 录

第 1 章 总论.....	1	第 3 章 统计整理	44
1.1 统计学概述.....	3	3.1 统计整理概述	45
1.1.1 统计学的含义.....	3	3.1.1 统计整理的重要性.....	45
1.1.2 统计学的产生与发展	3	3.1.2 统计整理的内容与程序.....	46
1.2 统计学的研究对象与研究方法	7	3.1.3 统计资料的审核	47
1.2.1 统计学的研究对象.....	7	3.1.4 统计资料的汇总技术.....	47
1.2.2 统计学的研究方法	7	3.1.5 统计汇总的组织形式.....	48
1.3 统计学的几个基本概念	9	3.2 统计分组	49
1.3.1 统计总体与总体单位	9	3.2.1 统计分组的意义	49
1.3.2 标志与统计指标	11	3.2.2 统计分组的作用	49
1.3.3 变异与变量.....	15	3.2.3 统计分组的要求与基本原则.....	51
1.3.4 统计指标体系.....	16	3.2.4 统计分组的方法	52
课后习题	17	3.2.5 统计分组的形式	52
第 2 章 统计调查.....	21	3.3 分配数列	54
2.1 统计调查的意义与种类.....	22	3.3.1 分配数列的概念与种类.....	54
2.1.1 统计调查的意义.....	22	3.3.2 变量数列的编制	55
2.1.2 统计调查的种类.....	24	3.4 统计表与统计图	59
2.2 统计调查方案.....	27	3.4.1 统计表	59
2.3 统计报表.....	30	3.4.2 统计图的绘制	61
2.3.1 统计报表的基本概念	30	课后习题	61
2.3.2 统计报表的种类.....	31	第 4 章 总量指标与相对指标	66
2.3.3 统计报表的资料来源.....	33	4.1 总量指标	67
2.4 统计专门调查.....	34	4.1.1 总量指标的概念与作用.....	67
2.4.1 普查.....	35	4.1.2 总量指标的种类	68
2.4.2 重点调查.....	37	4.1.3 总量指标的计算原则.....	72
2.4.3 典型调查.....	38	4.2 相对指标	72
2.4.4 抽样调查.....	39	4.2.1 相对指标的概念与表现	
课后习题	40	形式	72

4.2.2 相对指标的作用	73	课后习题	114
4.2.3 相对指标的种类及相应计算方法	74	第 6 章 时间数列	120
4.3 总量指标与相对指标的计算和运用	81	6.1 时间数列概述	121
4.3.1 计算和运用总量指标时应注意的问题	81	6.1.1 时间数列的含义与种类	121
4.3.2 计算和运用相对指标时应注意的问题	82	6.1.2 编制时间数列的原则	123
4.3.3 总量指标与相对指标相结合	82	6.2 时间数列的水平指标	125
课后习题	83	6.2.1 发展水平	125
第 5 章 平均指标与变异指标	90	6.2.2 平均发展水平	125
5.1 平均指标的概念与作用	92	6.2.3 增长量	128
5.1.1 平均指标的概念	92	6.2.4 平均增长量	129
5.1.2 平均指标的作用	93	6.3 时间数列的速度指标	130
5.2 平均指标的种类及相应计算方法	93	6.3.1 发展速度	130
5.2.1 算术平均数	93	6.3.2 增长速度	131
5.2.2 调和平均数	98	6.3.3 平均发展速度与平均增长速度	132
5.2.3 几何平均数	100	6.3.4 季节变动的测定	135
5.2.4 中位数	102	6.4 长期趋势的测定	139
5.2.5 众数	104	6.4.1 时距扩大法	139
5.2.6 平均指标间的相互关系	106	6.4.2 移动平均法	140
5.3 平均指标的应用原则	107	6.4.3 分割平均法	141
5.3.1 平均指标只能运用于同质总体	107	6.4.4 最小平方方法	143
5.3.2 用组平均数补充说明总平均数	107	6.4.5 拟合长期趋势方程时应考虑的问题	148
5.3.3 用分配数列补充说明总平均数	108	课后习题	149
5.4 标志变异指标	108	第 7 章 统计指数	156
5.4.1 标志变异指标的概念	108	7.1 统计指数概述	157
5.4.2 标志变异指标的作用	110	7.1.1 统计指数的意义	157
5.4.3 常用的标志变异指标	110	7.1.2 统计指数的作用	158
		7.1.3 统计指数的种类	159
		7.2 综合、区域与计划完成情况指数	160
		7.2.1 数量指标综合指数	162
		7.2.2 质量指标综合指数	163
		7.2.3 区域指数	165

7.2.4 计划完成情况指数	166	9.1.2 相关关系的主要特征与 种类	220
7.3 平均指数	167	9.1.3 相关分析的内容与作用	222
7.3.1 加权算术平均数指数	168	9.1.4 相关系数的测定	223
7.3.2 加权调和平均数指数	169	9.2 回归分析	225
7.3.3 平均数指数与综合指数的 比较	170	9.2.1 回归分析的特点	225
7.3.4 固定权数加权算术平均法	171	9.2.2 线性回归方程的确定	225
7.3.5 几种常用的经济指数	173	9.2.3 估计标准误差的测定	226
7.3.6 价格指数的应用	179	9.2.4 运用回归分析时应注意的 问题	228
7.4 指数体系与因素分析	180	9.2.5 多元线性回归分析	229
7.4.1 指数体系	180	课后习题	231
7.4.2 因素分析法概念	181	第 10 章 国民经济核算	236
7.4.3 总量指标的因素分析	182	10.1 国民经济核算概述	238
7.4.4 平均指标的因素分析	185	10.1.1 建立国民经济核算体系的 意义	238
课后习题	187	10.1.2 国民经济的两种核算 体系	238
第 8 章 抽样推断	194	10.1.3 国民经济产业划分	241
8.1 抽样推断概述	195	10.1.4 国民经济行业分类	242
8.1.1 抽样推断的概念与特点	195	10.1.5 经济类型分类	242
8.1.2 抽样推断的作用与内容	197	10.1.6 经济资产及分类	243
8.1.3 抽样推断的几个基本概念	198	10.2 国民经济核算的主要指标	244
8.2 抽样误差	201	10.2.1 经济流量与经济存量	244
8.2.1 抽样误差的概念	201	10.2.2 国内生产总值	245
8.2.2 影响抽样误差的因素	202	10.2.3 国内生产净值	247
8.2.3 抽样平均误差	202	10.2.4 国民总收入	247
8.2.4 抽样极限误差	204	10.2.5 国民净收入	247
8.2.5 抽样误差的概率度	205	10.2.6 国内生产总值增长率	247
8.3 抽样推断方法	207	课后习题	248
8.3.1 总体参数的点估计	207	附录 正态分析概率表	251
8.3.2 总体参数的区间估计	207	参考文献	253
课后习题	213		
第 9 章 相关与回归分析	218		
9.1 相关分析	219		
9.1.1 相关关系的概念	219		

第1章

总论



知识目标

掌握统计的含义和基本性质

掌握统计的基本概念及相互间关系



能力目标

能正确认识统计对宏观与微观经济的作用，提高对社会经济形势的认识能力

2035

2547





开篇案例

统计学的威力

汰渍是宝洁公司的一个洗衣粉品牌，奥妙是联合利华公司的洗衣粉品牌，双方都曾在电视上大打广告。然而有一段时间，宝洁公司发现它的汰渍洗衣粉销量不如奥妙。分析人员搜集了消费者对于广告效应的反映数据，经过分析后发现有一个很细小的环节——电视画面上把汰渍洗衣粉往洗衣机里倒的时间要比奥妙长一点点，因此带来一种暗示——汰渍的洗涤效果赶不上奥妙。找到原因后，宝洁公司立刻调整了电视广告，使得产品销量继续增长起来。

二战后，日本经济迅速崛起，其中的一个重要原因是产品质量大大提高，而质量提高在相当程度上是得益于统计思想和方法的应用，爱德华兹·戴明就是当时在日本传授统计质量管理思想的先驱之一，日本人接受并发展了他的思想。值得讽刺的是，作为美国联邦统计局官员的戴明，直到1980年众多日本产品占领美国市场时才被美国人认识。此时，时间已过去了30年。

几个世纪以来，学者们怀疑是不是莎士比亚写下那些精彩绝伦的戏剧和诗句，怀疑的部分原因是人们认为莎士比亚这个历史人物是乡间的、没文化的，而且未发现他取得如此巨大的成功。他在世时，没有人认为他是作家。他死后，没留下任何私人信件、手稿或文学笔记。因此人们怀疑威廉·莎士比亚可能另有其人。一些著名的、受过良好教育的、有着丰富的文学记录的人，被认为写下了这些经典作品，他们中有 Francis Bacon, John Donne 等。

一些专家利用统计侦探方法帮助莎士比亚得到了他的应得之物，这种统计方法依赖于众数的使用。

在3年时间里，一个由 Claremont 学院本科学生组成的莎士比亚诊所，用统计分析对58个与莎士比亚同时代的作家进行分析，以确定谁的写作风格与莎士比亚作品的风格最接近。他们从58个作家的作品中抽取片段，并将其分成500字一段的小段。他们对分组中一些变量进行技术统计。例如，学生们考察52个关键字出现的情况，并找出众数，利用各种统计策略，他们得到了各个作家的主要特征，调查结果如下。

27个被选者的诗中，没有一个能通过众数检验。Thomas Heywood 是最接近的一个检验，他和莎士比亚相差2.2个标准差。这意味着，如果 Heywood 真是莎士比亚，那么和他以自己的名字命名的诗篇相同的机会将会少于5%。John Donne 是最远的检验，与莎士比亚作品的标准差达到36.9。这些明显的事实证明，就是莎士比亚写下了他本人的诗篇。

西单附近有一家餐厅叫文府大厨，饭菜的品质非常不错。但因为种种原因，客源并不好，老板抱着试试看的态度成了民生银行的特约商户。该银行在持卡人数据库里根据消费

者在餐厅的消费记录,优选了餐厅消费记录特别多的消费者,把文府大厨的菜品和打折信息介绍给他们。结果,就在当周的周末,该银行信用卡部门的人员于中午到这家餐厅进行考察时,所有的座位都坐满了,其中 75% 的顾客是该银行的持卡人。银行数据库的建立,有价值信息的提取,都和统计学有一点关系,确切地说这是统计学同计算机科学的一个交叉分支——数据挖掘的典型应用。

资料来源:吕朝辉. 统计学基础. 北京: 化学工业出版社. 2008.

1.1 统计学概述

1.1.1 统计学的含义

“统计”一词是由英语 statistics 翻译过来的,源于古希腊语 state(城邦)、拉丁语 status(状态)以及意大利语 stato(国家)和 statisti(政治家)。可见,历史上统计一词的含义与国家的政治经济活动有关,它是用来记载和描述一个国家重要政治经济事项的方法。在汉语中,“统计”是对大量事物的数量方面进行的汇总和计数,即“统而计之”的意思。随着社会经济的发展,人们对统计的功能提出了更高的要求,统计一词的内涵和外延都发生了巨大的变化。从不同的角度理解,统计一词具有 3 种不同的含义:统计工作、统计资料和统计学。

统计工作是人们对客观事物数量方面进行调查研究的认识活动,包括数据资料的搜集、整理和分析等。统计工作是统计一词最基本的含义,没有统计工作就不会生成统计资料,而没有丰富的统计实践活动,也不会产生统计科学。

统计资料是统计工作的成果,包括调查得到的、经过整理的、具有信息价值的各种统计数据和分析报告。

统计学则是一门关于研究客观事物数量方面和数量关系的方法论科学。统计学是统计工作实践经验的理论概括,同时,它又反过来从理论和方法上指导统计工作,推动统计工作不断发展。两者是理论与实践的辩证统一的关系。

1.1.2 统计学的产生与发展

1. 统计学的产生

统计学是随着社会生产发展和适应国家管理的需要而产生和发展起来。在原始社会里,人类最初的一般计数活动,蕴藏着统计的萌芽。随着奴隶制国家的产生,统治阶级为了对内统治和对外战争,需要征兵征税,开始了对人口、土地和财产的统计。中国从公元前 2000 多年的夏朝,就有了人口、土地的统计数字记载。例如,分中国为九州,人口 1 355 万,土地约 2 438 万顷。在差不多同一时期的古希腊、罗马等奴隶制国家里,也有了人口、财产和世袭领地的统计。

2. 统计学的发展

在整个奴隶社会和封建社会时期,统计已略具规模。封建君主和精明的政治家日益意识到统计对于治国强邦的重要作用。例如,我国战国时期的商鞅就提出,要使国家强盛就必须知道“境内仓、口府之数,壮男、壮女之数,老、弱之数,官士之数,以言说取食者之数,利民之数,马、牛、刍、藁之数”等所谓的“十三数”;至秦汉,已有地方田亩和户口的记录;唐宋则有计口授田、田亩鱼鳞册等土地调查和计算;到明清,又建立了经常性的人口登记和保甲制度等。到了封建社会末期,特别是进入资本主义社会以后,社会生产力迅速发展,社会分工愈益精细,交通、航运、贸易日趋发达,国际市场逐步形成。事实上,不仅政府需要有包括人口、土地、财富、赋税和军事等方面国情国力的统计,各类企业主、商人为了经营管理和争夺市场,也需要各种商业情报和市场信息,统计逐步扩展到工业、农业、贸易、银行、保险、交通、邮电、海关等部门,并且出现了专业的统计机构和研究组织。统计逐步成为社会分工中的一个独立部门和专业。

从17世纪下半叶开始,欧洲出现了一些统计理论著作,并逐步形成了不同的学派。从统计学的产生和发展来看,可以把统计学大致分为古典统计学、近代统计学和现代统计学3个时期。

1) 古典统计学时期

这是指17世纪中叶至18世纪中叶统计学萌芽时期。当时,有记述学派和政治算术学派两大学派。

(1) 记述学派。记述学派又称国势学派,产生于18世纪封建制度的德国。所谓国势学就是以文字记述国家显著事项的学说。由于当时德国许多大学设有国势学这门课程,故国势学派亦称德意志大学教授派。其创始人康令(H. Coning, 1606~1681),主要代表人物为阿亨瓦尔(G. A. Chenwall, 1719~1772)。最早讲授国势学的是康令,他第一个在德国赫尔莫斯达德大学讲授“欧洲最近国势学”,奠定了国势学的基础。阿亨瓦尔在哥丁根大学开设“国势学”课程,其主要著作是《近代欧洲各国国势学概论》,书中讲述“一国或多数国家的显著事项”,主要是用对比分析的方法,研究关于国家组织、人口、军队、领土、财产等国情国力,比较各国实力的强弱,为德国的君主政体服务。这个学派在进行国势比较分析中,偏重事物性质的解释,而不注重数量分析。实际上,国势学派研究的是历史学的组成部分,属于实质性的社会科学。这个学派被统计史学家称为有其名而无其实,因为“统计学”学科的名词,便是由阿亨瓦尔在1794年首先提出的;此外,这个学派还广泛应用了“统计数字资料”、“数字对比”等统计术语和对比的方法,这些术语和方法为后人所继承,沿用至今。

随着资本主义市场经济的发展,对事物量的计算和分析显得越来越重要。该学派后来发生了分裂,分化出表式学派,并逐渐发展为政府统计。

(2) 政治算术学派。政治算术学派产生于 17 世纪中叶的资本主义国家英国,代表人物是威廉·配第(W.Petty, 1623~1687),其代表作是《政治算术》一书。在这部著作中,他以数字资料为基础,用计算和对比的方法对英、法、荷 3 国的实力进行了比较,论证了英国称雄世界的条件和地位。尤为重要的是,他所采用的独特的方法,即“对于人口、土地、资本、产业的真实情况的表述方法”。他主张一切论述都用数字、重量和尺度来进行,只重视“诉诸感觉的”,即经验上的论证,反对受主观因素左右的思辨的议论。他还提出了用图表形式概括数字资料的理论和方法。这种理论和方法对后来统计学的形成和发展有深远的影响。马克思称他为“政治经济学之父”,在某种程度上也可以说他是统计学的创始人。

政治算术学派用计量方法,即大量观察法、分类法以及对比法,综合研究社会经济问题,具有开创性的意义。尽管当时还未利用统计学之名,却已有统计学之实了。

2) 近代统计学时期

18 世纪末到 19 世纪末的 100 多年中,统计学有了很大发展,又形成了许多学派,其中主要是数理统计学派和社会统计学派。

(1) 数理统计学派。数理统计学派产生于 19 世纪中叶,它是在概率论已有相当发展的基础上,把概率论引进统计学而形成的。其奠基人是比利时物理学家和统计学家凯特勒(A.Quetelet, 1796~1874),主要代表作有《社会物理学》等。

凯特勒师承法国数字学家、统计学家拉普拉斯(P.s.LaPlace, 1794~1827),主张用研究自然科学的方法研究社会现象。他正式把概率论引进统计学,使统计学进入一个新的阶段;他最先用大数定律论证了社会生活现象纷繁复杂、变化不定的偶然性中存在着规律性,并且提出了误差理论,用来解决统计上的准确性问题。在学科性质问题上,他明确地认为统计学是一门既研究社会现象又研究自然现象的独立的方法论科学。饶有趣味的是,他运用自己的理论和方法提出了“平均人”的概念,塑造了一个具有平均身高、平均体重、平均智力和道德品质的人物“模特”,宣称统计的任务就是关于“平均人”的比较研究。如果社会上所有的人同平均人没有差异,社会矛盾就会趋于缓和。

凯特勒混淆自然现象和社会现象的本质界限,用自然规律来解释诸如犯罪、道德等社会问题,做出一些机械论的、庸俗化的解释,这是他的历史局限性。但是,他把概率论同统计结合起来,却为后来数理统计学的形成与发展奠定了基础。

随着社会生产的发展和自然科学技术的进步,统计研究的领域不断扩大,概率论方法的运用也日益增多并渐趋成熟。到 19 世纪末,数理统计学就从统计学中分离出来自成一派。由于它主要是在英、美等国发展起来的,故又称英美数理统计学派。

(2) 社会统计学派。19 世纪后半叶,正当英美数理统计学派开始发展的时候,在欧洲又兴起了社会统计学派。社会统计学派的主要代表人物有恩格尔(C.L.E.Engel, 1821~1896)以及梅尔(G.v.Mayr, 1841~1925)等。他们认为统计学是一门社会科学,是研究社会现象变动原因和规律性的实质性科学,以此同数理统计学派的通用方法论相对立。