

教育部经济管理类核心课程教材

国家级精品课程教材



统计学

曾五一 朱建平 主编

STATISTICS
STATISTICS

上海财经大学出版社

教育部经济管理类核心课程教材

国家级精品课程教材



统计学

曾五一 朱建平 主编

STATISTICS
STATISTICS

上海财经大学出版社

图书在版编目(CIP)数据

统计学/曾五一,朱建平主编. —上海:上海财经大学出版社,2012.8

(教育部经济管理类核心课程教材)

ISBN 978-7-5642-1357-2/F·1357

I. ①统… II. ①曾… ②朱… III. ①统计学-高等学校-教材

IV. ①C8

中国版本图书馆CIP数据核字(2012)第092305号

- ☐ 责任编辑 张美芳
- ☐ 书籍设计 钱宇辰
- ☐ 责任校对 赵伟 卓妍

TONGJIXUE

统 计 学

曾五一 朱建平 主编

上海财经大学出版社出版发行
(上海市武东路321号乙 邮编200434)

网 址: <http://www.sufep.com>

电子邮箱: webmaster@sufep.com

全国新华书店经销

上海市印刷七厂印刷

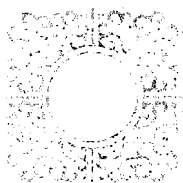
上海远大印务发展有限公司装订

2012年8月第1版 2012年8月第1次印刷

787mm×1092mm 1/16 19.75印张 512千字

印数:0 001—5 000 定价:39.00元

前 言



统计学是研究如何测定、收集、整理、归纳和分析反映客观现象总体数量的数据,以帮助人们正确认识客观世界数量规律的方法论科学。现代统计学已经渗透到理、工、农、医、经济管理与各类人文社会科学领域,并由此产生许多新的交叉学科,改变了原有单一学科发展的思路,对各门科学的发展具有极大的支撑作用;反过来,这种结合又促进了统计学自身的发展。因此,统计学被称为 21 世纪最有发展前途的学科之一。长期以来,统计学一直被教育部列为不少专业的核心课程。为了让更多的社会、经济和管理工作者更好地领会统计思想,灵活地应用统计方法来认识客观世界数量规律,并指导实际工作,应上海财经大学出版社的邀请,我们组织编写了这本适用于高等院校财经管理类各专业以及其他社科类专业如社会学、法学、新闻学、政治学等专业的统计学教材。

本教材的编写力求以统计思想为主线,深入浅出地介绍各种统计方法的应用。

本教材的特点是:

第一,努力贯彻“少而精”和“学以致用”的原则。将原有的数理统计和社会经济统计原理两门课程的内容有机地结合起来。

第二,注重统计思想的讲述。从实际问题入手,淡化统计方法本身的数学推导,体现统计学的实用性。

第三,应用 Excel 软件,实现统计计算。选用最常见的通用软件 Excel 作为计算工具,在附录中专门介绍 Excel 的应用,并与教材内容相结合实现统计计算。

第四,根据教学的需要,在出版教材的同时,编制教学用的幻灯片和其他教学辅导材料。

本教材各章初稿的执笔人是:曾五一(第一、八章)、朱建平(第五、六、七章)、黄长全(第三、四章)、庄贇(第九〈合作〉、十一章〈合作〉)、方匡南(第二、九章〈合作〉)、许永洪(第十、十一章〈合作〉)、袁加军(附录一)。全书由曾五一、朱建平担任主编,负责整体设计、修改、审订、总纂与定稿工作。

编写一本好的教材并不容易,尽管我们努力想奉献给读者一本满意的教材,但书中仍难免有疏漏甚至错误之处,恳请读者多提宝贵意见。

编者

2012 年 2 月

目 录

前言	(1)
第一章 绪论	(1)
第一节 什么是统计	(1)
一、统计的涵义	(1)
二、统计研究的基本环节	(2)
第二节 统计学的产生与发展	(4)
一、统计学的主要流派	(4)
二、统计学发展的新动向	(5)
三、理论统计学和应用统计学	(6)
四、统计学与有关学科的联系与区别	(7)
第三节 统计学的基本概念	(9)
一、总体与总体单位	(9)
二、样本	(10)
三、标志	(10)
四、统计指标与指标体系	(10)
思考与练习	(11)
第二章 数据的收集、整理与显示	(13)
第一节 数据的收集	(13)
一、数据概述	(13)
二、数据收集的方法及形式	(16)
三、统计调查体系及方案设计	(17)
四、间接统计数据的主要来源	(22)
第二节 数据的整理	(23)
一、数据整理概述	(23)
二、统计分组	(23)
三、频数分布	(28)
第三节 数据的显示	(32)
一、统计表	(32)
二、统计图	(34)

三、统计分析报告	(37)
思考与练习	(38)
第三章 数据分布特征描述	(40)
第一节 统计变量集中趋势的测定	(40)
一、测定集中趋势的意义	(40)
二、位置代表值	(40)
三、数值平均数	(43)
第二节 统计变量离散程度的测定	(53)
一、测定离散程度的意义	(53)
二、极差、四分位差和平均差	(54)
三、方差与标准差	(56)
四、离散系数与异众比率	(60)
第三节 变量分布的特征描述	(61)
一、矩的概念	(61)
二、偏度	(62)
三、峰度	(63)
思考与练习	(65)
第四章 概率基础	(68)
第一节 概率的基本概念	(68)
一、随机试验与随机事件	(68)
二、概率	(69)
第二节 随机变量及其分布	(71)
一、随机变量与概率分布的概念	(71)
二、概率分布的类型	(72)
三、随机变量的数字特征	(76)
四、随机向量与独立性	(78)
第三节 几种常见的概率分布	(81)
一、离散型分布	(81)
二、连续型分布	(81)
第四节 大数定律与中心极限定理	(83)
一、大数定律	(83)
二、中心极限定理	(84)
思考与练习	(85)
第五章 参数估计	(87)
第一节 抽样分布	(87)
一、抽样的基本概念	(87)
二、抽样分布	(88)

第二节 点估计与估计量的评价标准	(91)
一、点估计	(91)
二、估计量的优良标准	(93)
第三节 简单随机抽样的区间估计	(95)
一、总体均值的置信区间	(95)
二、总体成数的置信区间	(97)
三、两个总体均值及两个总体成数之差的置信区间	(98)
四、样本容量的确定	(101)
第四节 复杂随机抽样的区间估计	(103)
一、分层抽样的估计	(103)
二、等距抽样的估计	(104)
三、整群抽样的估计	(106)
四、多阶段抽样的估计	(108)
思考与练习	(109)
第六章 假设检验	(113)
第一节 假设检验的基本原理	(113)
一、假设检验的基本原理	(113)
二、假设检验的规则与两类错误	(114)
三、检验功效	(116)
第二节 总体参数假设检验	(117)
一、总体均值的假设检验	(117)
二、两个总体均值之差的检验	(120)
三、总体成数的假设检验	(121)
四、正态总体方差的假设检验	(123)
五、两个正态总体方差比的检验	(124)
第三节 非参数检验	(126)
一、非参数检验概述	(126)
二、 χ^2 检验	(126)
三、符号检验	(128)
四、秩和检验	(131)
五、游程检验	(132)
六、等级相关	(133)
思考与练习	(134)
第七章 方差分析	(138)
第一节 方差分析方法引导	(138)
一、方差分析问题的提出	(138)
二、方差分析的基本原理	(139)
第二节 单因素方差分析	(140)

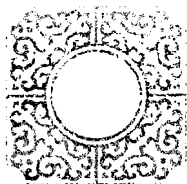
一、单因素条件下的平方和分解公式	(140)
二、因素作用显著性的检验	(141)
三、应注意的问题	(143)
第三节 双因素方差分析	(143)
一、无交互作用的双因素方差分析	(144)
二、有交互作用的双因素方差分析	(147)
思考与练习	(150)
第八章 相关与回归分析	(154)
第一节 相关与回归分析的基本概念	(154)
一、函数关系与相关关系	(154)
二、相关关系的种类	(154)
三、相关分析与回归分析	(155)
四、相关图	(156)
第二节 简单线性相关分析	(156)
一、相关系数及其检验	(156)
二、等级相关系数的计算和检验	(159)
第三节 一元线性回归分析	(160)
一、标准的一元线性回归模型	(160)
二、一元线性回归模型的估计	(162)
三、一元线性回归模型的检验	(166)
四、一元线性回归模型预测	(168)
第四节 多元线性相关与回归分析	(172)
一、标准的多元线性回归模型	(172)
二、多元线性回归模型的估计	(172)
三、多元线性回归模型的检验	(174)
四、多元线性回归预测	(176)
五、复相关系数和偏相关系数	(179)
第五节 非线性相关与回归分析	(181)
一、非线性函数形式的确定	(181)
二、非线性回归模型估计	(183)
三、相关指数	(184)
思考与练习	(185)
第九章 时间序列分析	(189)
第一节 时间序列分析概述	(189)
一、时间序列的概念	(189)
二、时间序列的种类	(190)
三、时间序列的编制方法	(191)
第二节 时间序列的分析指标	(192)

一、时间序列分析的水平指标	(192)
二、时间序列分析的速度指标	(198)
三、水平分析与速度分析的结合与应用	(202)
第三节 长期趋势的测定	(203)
一、时间序列的构成与分解	(203)
二、长期趋势的测定方法	(204)
第四节 季节变动和循环波动测定	(209)
一、季节变动的测定方法	(209)
二、循环变动的测定方法	(212)
三、随机变动的测定方法	(213)
第五节 时间序列预测方法	(214)
一、趋势外推法	(214)
二、自回归预测法	(214)
三、移动平均预测法和指数平滑法	(215)
思考与练习	(220)
第十章 对比分析与指数分析	(223)
第一节 对比分析	(223)
一、对比分析的概念	(223)
二、对比分析指标的计算	(223)
第二节 指数的概念和种类	(225)
一、统计指数的概念	(225)
二、统计指数的种类	(225)
三、统计指数的作用	(227)
第三节 综合指数	(227)
一、综合指数的编制原理	(227)
二、拉氏指数和帕氏指数	(228)
三、其他形式的综合指数	(230)
第四节 平均指数	(231)
一、平均指数的编制原理	(231)
二、加权算术平均数指数	(231)
三、加权调和平均数指数	(232)
第五节 指数因素分析	(233)
一、指数体系的概念	(233)
二、两因素指数分析	(233)
三、多因素指数分析	(236)
第六节 几种常见的经济指数	(238)
一、工业生产指数	(238)
二、物价指数	(239)
三、股票价格指数	(241)

思考与练习	(242)
第十一章 统计综合评价	(244)
第一节 统计综合评价概述	(244)
一、统计综合评价的意义	(244)
二、统计综合评价的基本步骤	(244)
第二节 评价指标选择与数据预处理	(246)
一、评价指标的选择	(246)
二、数据预处理方法	(247)
第三节 权重的确定与评价结果的综合	(250)
一、确定权重的基本方法	(250)
二、评价结果的综合	(253)
思考与练习	(256)
附录一 Excel 在统计中的应用	(257)
第一节 Excel 简介与基本操作	(257)
一、Excel 简介	(257)
二、熟悉 Excel 工作界面	(257)
三、输入数据	(262)
四、编辑工作表	(263)
五、使用公式和函数	(265)
第二节 分析工具库与统计函数	(271)
一、分析工具	(271)
二、常用数学与统计函数简表	(272)
第三节 Excel 在描述统计中的应用	(274)
一、编制分布数列	(274)
二、绘制统计图	(276)
三、计算描述统计指标	(278)
第四节 Excel 在概率计算和统计推断中的应用	(280)
一、概率计算	(280)
二、区间估计	(281)
三、假设检验	(282)
四、方差分析	(286)
第五节 Excel 在回归分析中的应用	(288)
第六节 Excel 在时间序列分析中的应用	(291)
一、求解按方程法计算的平均发展速度	(291)
二、长期趋势的测定	(293)
三、季节变动与循环变动的测定	(293)
四、时间序列模型预测	(294)

附录二 常用统计表	(296)
附表 2-1 正态分布概率表	(296)
附表 2-2 t 分布临界值表	(299)
附表 2-3 χ^2 分布临界值表	(300)
附表 2-4 F 分布临界值表	(301)
附表 2-5 二项分布临界值表	(303)
附表 2-6 秩和检验表	(304)
附表 2-7 游程检验 R 临界值表	(305)
附表 2-8 斯皮尔曼等级相关系数 r_s 的上临界值 r_α 表	(306)

第一章



绪 论

第一节 什么是统计

一、统计的涵义

以下是关于“统计”的几则信息：

我国著名经济学家马寅初指出：“学者不能离开统计而治学，政治家不能离开统计而施政，事业家不能离开统计而执业。”

美国著名科幻作家威尔斯说：“对于追求效率的公民而言，统计思维总有一天会和读写能力一样重要。”

根据媒体的报道，美国总统的年薪在各国元首中名列前茅，但根据美国 2004 年《工作等级年鉴》一书的排名，总统一职并未进入最好工作之列，在美国最好的工作是一——统计学家。

由以上几则信息可知，统计已经渗透到社会经济活动和科学研究的方方面面，统计无处不在，并且发挥着越来越重要的作用。那么，究竟什么是统计？统计是如何开展研究的？作为一门科学的统计学与其他学科有何区别与联系？这些正是本章所要介绍的主要内容。

在日常生活中，人们对于“统计”这一术语常常有不同的用法。例如，企业每年要“统计”产量和产值，这是将其作为一种工作来看待。了解股票的交易状况要看有关成交额和股票指数“统计”，这时又是将其作为数据来运用。而大学课堂上所说的我们正在学习“统计”，则是指一门科学，即统计学。

那么究竟何为统计，这里有必要给出一个比较准确的科学定义。所谓统计，它是人们认识客观世界总体数量变动关系和变动规律的活动的总称，是人们认识客观世界的一种有力工具。统计的研究对象具有以下特点：

（一）数量性

这是统计研究对象的基本特点，常言说：“数字是统计的语言”，“数据是统计的成果”，指的正是这个意思。但并不是任何一种数量都可以作为统计对象。统计数据是客观事物量的反映，统计定量认识必须建立在对客观事物定性认识的基础上。

（二）总体性

统计的数量研究是对现象总体中各单位普遍存在的事实进行大量观察和综合分析，得出反映现象总体的数量特征。例如，进行城镇居民家计调查，需要对具体的居民家庭进行调查，

但其目的并不在于了解个别居民家庭的生活状况,而是要反映一个国家、一个城市的居民收入水平、收入分配、消费水平、消费结构等。

(三) 变异性

统计研究同类现象总体的数量特征,它的前提是总体各单位的特征表现存在着差异,而且这些差异并不是事先可以预知的。例如,各种股票的价格和成交量每天不同,这才需要对其进行统计,编制股票指数等指标。如果说总体各单位的变异表现出个别现象的特殊性和偶然性,那么对现象总体的数量研究则是通过大量观察从各单位的变异中归纳概括出它们的共同特征,显示出现象的普遍性和必然性。

统计作为一种社会实践活动有悠久的历史。据历史记载,我国在西周朝代就已建立了统计报告制度。在英文中,统计为 statistics,它与“国家”为同一词根。可以说,自从有了国家,就有统计实践活动。最初,统计只是为统治者了解国家的情况和管理国家提供数量依据。随着社会经济和科技的发展以及统计学自身的进步,统计的应用领域不断扩大。现在,统计不仅被用于经济管理领域,而且在医学、生物、物理等其他领域也得到广泛的应用。

人们通过统计实践活动所得的数据即**统计数据**,它既是统计工作的成果,也是进一步开展统计研究的基础。人们对统计数据的通常要求是:客观性,即它能反映客观事实而不受任何偏见的影响或任何势力的干扰;适用性,即统计数据能够适应统计研究的目的;准确性,即统计数据能够反映真实的情况,不出现较大的误差;及时性,即统计数据应及时收集、及时加工、及时公布。

统计学是在统计实践的基础上,产生并逐步发展起来的一门科学。统计学与统计实践活动的关系是理论与实践的关系,理论源于实践,又高于实践,反过来又指导实践。

根据统计学界当前的主流观点,可以给出以下关于**统计学的基本定义**。“统计学是研究如何测定、收集、整理、归纳和分析反映客观现象总体数量的数据,以帮助人们正确认识客观世界数量规律的方法论科学。”根据上述基本定义,可以对统计学的学科性质和研究内容有一个比较清楚的认识。首先,统计学研究的目的是帮助人们更好地认识客观世界的数量规律。其次,统计学是一门方法论的科学,其主要内容并不是对客观世界的数量规律本身及其产生的机理进行研究,而是为这一研究提供合适的方法和工具。最后,统计学的研究是围绕反映客观现象总体数量的数据展开的。统计研究的重要特点是“用数据说话”,正是在这个意义上,统计学又被称为“数据科学”。

二、统计研究的基本环节

利用统计认识客观世界数量规律的全过程包括以下几个基本环节:

(一) 统计设计

根据所要研究问题的性质,在有关学科理论的指导下,制定统计指标、指标体系和统计分类,给出统一的定义、标准,同时提出收集、整理和分析数据的方案和工作进度等。统计设计是整个统计研究的前期工程,其完成质量直接关系到整个统计研究的质量。搞好统计设计不仅要有统计学的一般理论和方法为指导,而且还要求设计者对所研究的问题本身具有深刻的认识和相关的学科知识。例如,要设计一套较好地评价企业经营状况的统计体系与方案,仅有一般的统计方法知识是不够的,设计者还必须具备企业经营管理知识和理论素养。

(二) 收集数据

经过统计设计形成方案之后,就可以开始收集统计数据。统计数据的收集有两种基本方

法。对于大多数自然科学和工程技术研究来说,有可能通过有控制的科学实验去取得数据,这时可以采用实验法。在统计学中有专门的一个分支——实验设计,就是研究如何科学地设计实验方案,从而使得通过实验采集的数据能够符合分析的目的和要求。对于社会经济现象来说,一般无法进行重复实验,要取得有关数据就必须到社会总体中去选取足够多的单位进行调查观察,并加以综合研究。如何科学地进行调查是统计学研究的重要内容。本书是为经济与管理类专业编写的统计学入门教科书,由于篇幅的限制,将只介绍有关统计调查的理论与方法。

(三)整理与分析

原始的统计数据收集上来之后,还必须经过整理、加工和分析才能真正发挥其作用。在统计研究的这一阶段,所运用的方法包括两大类:描述统计和推断统计。

描述统计是指对采集的数据进行登记、审核、整理、归类,在此基础上进一步计算出各种能反映总体数量特征的综合指标,并用图表的形式表示经过归纳分析而得到的各种有用的统计信息。统计描述是统计研究的基础,它为统计推断、统计咨询、统计决策提供必要的事实依据。统计描述也是对客观事物认识的不断深化过程。它先对分散无序的原始数据进行整理归纳,再运用分组法和综合指标法得到现象总体的数量特征,揭示客观事物内在的数量规律性,以达到认识的目的。

推断统计是在对样本数据进行描述的基础上,利用一定的方法根据样本数据去估计或检验总体的数量特征。在进行统计研究时,常常存在这种情况,由于各种原因,我们所掌握的数据只是部分单位的数据或有限单位的数据,而我们所关心的却是整个总体的数量特征。例如,民意测验中某一候选人能否当选?全国婴儿的性别比例如何?某种电子元件的寿命多长?这时就必须利用统计推断的方法来解决。统计推断是逻辑归纳法在统计推理中的应用,所以又称归纳推理。推断统计是现代统计学的主要内容。

(四)统计资料的积累、开发与应用

通过统计整理和分析,可以得到有关的统计资料,但统计资料的提供并不意味着统计研究的终结。统计的目的在于认识客观世界的数量规律,仅凭一次收集的统计资料,往往还不能很好地发现客观世界存在的数量规律。因此,对于已经公布的统计资料需要加以积累,同时还可以进一步加工,利用各种数量模型并结合相关的实质性学科的理论知识进行分析。如何更好地将统计资料和统计方法应用于各自的研究领域,是应用统计学研究的一个重要方面。

以上所述的统计研究的全过程可以用图 1-1 表示:

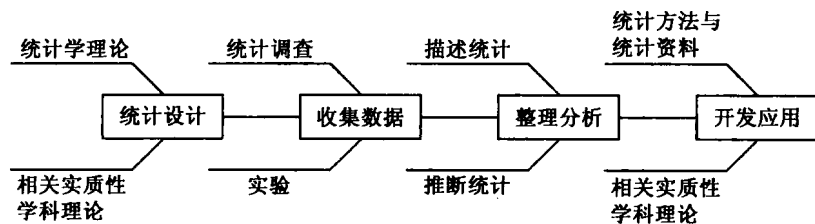


图 1-1 统计研究的全过程

第二节 统计学的产生与发展

一、统计学的主要流派

虽然人类统计的实践活动可以追溯到相当遥远的古代,但是,将统计实践上升到理论,并总结和概括成一门科学——统计学,距今只有 300 多年的历史。从统计学发展的历史来看,曾经产生过较大影响的主要有以下几个流派:

(一)政治算术学派

最早的统计学源于 17 世纪的英国,其代表人物是威廉·配第(William Petty, 1623~1687)和约翰·格朗特(John Graunt, 1620~1674)。威廉·配第在其代表作《政治算术》(1676 年)一书中写道:本书“不用比较级、最高级进行思辨或议论,而是用数字……来表达自己的问题……借以考察在自然中有可见的根据的原因”。约翰·格朗特在《对死亡表的自然观察和政治观察》(1662 年)一书中,通过大量观察的方法,研究并发现了一系列关于人口的数量规律,并运用各种方法对统计数据进行间接的推算和相互印证。威廉·配第等人关于运用大量观察和数量分析等方法对社会经济现象进行研究的主张,为统计学的发展开辟了广阔的前景。有趣的是,这一派学者一直没有使用“统计学”这一术语,而是用“政治算术”来表明其研究的特色,因而被称为政治算术学派。

(二)国势学派

最早使用“统计学”这一术语的是德国的阿亨瓦尔(G. Achenwall, 1719~1772)。他认为统计学是关于国家显著事项的学问,主要通过对国家组织、人口、军队、领土、居民职业以及资源财产等事项的记述对国情国力进行研究。后人把从事这方面研究的德国学者称为“国势学派”。国势学派虽然创造了“统计学”这一术语,但他们主要使用文字记述的方法进行研究,其学科内容与现代统计学有较大的差别。

(三)社会统计学派

19 世纪以后,随着经济和社会的发展,统计在社会经济领域中的应用越来越广泛、越来越深入。为满足国家和社会需要,人们广泛地开展各种统计调查活动。这不仅为经济学家和社会学家的理论研究和实证分析提供了数量依据,也为统计学家从中概括和提出新的统计方法提供了新思路和数据材料。包括政治统计、人口统计、经济统计、犯罪统计、社会统计等多方面内容的“社会统计学”一词开始出现,并成为统计学中的一个重要流派。该学派在德国、日本和苏联都有相当大的影响。1850 年,德国统计学家克尼斯(K. G. A. Knies)发表了题为《独立科学的统计学》的论文,提出统计学是一门独立的社会科学,是一门对社会经济现象进行数量对比分析的科学,他主张以“国家论”作为国势学的科学命名,而以“统计学”作为“政治算术”的科学命名。这一主张得到当时大多数统计学家的赞同。于是,以往无“统计学”之名而有“统计学”之实的“政治算术”取代了过去有“统计学”之名而无“统计学”之实的“国势学”成为统计学的正统。

社会统计学派着重对社会经济领域的统计方法及其应用进行研究。各国学者在社会经济统计指标的设定与计算、指数的编制、资料的收集与整理、统计调查的组织和实施、经济社会的数量分析和预测等方面做出的贡献已成为现代统计学的重要组成部分。例如,德国统计学家恩格尔(Christian Lorenz Ernest Engel, 1821~1896)提出的“恩格尔系数”,至今仍为人们广泛使用。由美国经济学家库兹涅茨(Simon Kuznets, 1901~1985)和英国经济学家斯通(Richard Stone, 1913~

1991)等人开发的国民收入和国内生产总值的核算方法,被称为“20 世纪最伟大的发明之一”。

(四)数理统计学派

数理统计学的创始人是比利时统计学家凯特勒(Adolphe Quetelet, 1796~1874),他的代表作《概率论书简》、《社会物理学》等将概率论和统计方法引入社会经济方面的研究。在学科性质上,凯特勒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。此后,以概率论为基础建立的统计理论与方法被称为数理统计。从 19 世纪中叶到 20 世纪中叶,数理统计学得到迅速发展:英国生物学家高尔顿(Fancis Galton, 1822~1911)首次提出并阐述了“相关”的概念;英国统计学家皮尔逊(Karl Pearson, 1857~1936)提出了计算复相关和偏相关的方法;英国统计学家戈塞特(William Sealy Gosset, 1876~1937)建立了“小样本理论”,即所谓的“ t 分布”;英国统计学家费歇(Ronald Aylmer Fisher, 1890~1962)在样本相关系数的分布、方差分析、实验设计等方面的研究中做出了重要贡献。到 20 世纪中期,数理统计学的基本框架已经形成。随着社会经济的发展和自然科学技术的进步,统计研究的领域不断扩大,数理统计方法的运用也日益广泛和深入,数理统计学派成为英美等国统计学界的主流。

在数理统计学派内部,围绕着对概率的不同认识,又可分为“经典学派”和“贝叶斯学派”。“经典学派”以实际观测的频率作为测定概率的基础,开展各种统计推断。“贝叶斯学派”则提出综合利用先验信息和后验信息来确定概率,并在此基础上进行统计推断的思路和方法。“贝叶斯学派”的代表人物之一沃尔德(Abraham Wald)在统计决策理论领域做出了重要贡献。

二、统计学发展的新动向

自 20 世纪 80 年代以来,随着社会经济的发展和科学技术的进步,统计学的发展又出现了一些新的动向。

首先,作为方法论科学的统计学与各实质性学科的结合越来越紧密。统计学已渗透到许多领域,并由此产生不少新的学科,改变了原有单一学科发展的思路,对其他科学的发展具有极大的支撑作用,反过来,这种结合又促进了统计学的发展。例如,统计学与经济学的结合产生了计量经济学这一新的学科。在诺贝尔经济学获奖者中,2/3 以上的研究成果与统计和经济定量分析有关。因此,萨缪尔森在其经典的教科书《经济学》(第 10 版)中特别提到:“在许多与经济学有关的其他学科中,统计学是特别重要的。”再如,近年来在发达国家兴起并取得很好效益的“6 西格玛管理”本质上就是将统计思想、统计方法和统计数据以及企业管理的理论结合起来,应用于企业的经营管理与质量控制。随着统计学的发展,在不少原来似乎很难与统计挂钩的领域也开始引进统计学的思想和研究方法。1981 年,首届国际《红楼梦》研讨会在美国召开,威斯康星大学的学者陈炳藻独树一帜,宣读了题为《从词汇上的统计论〈红楼梦〉作者的问题》的论文。他从字、词出现的频率入手,通过计算机进行统计、处理、分析,对《红楼梦》后 40 回系高鹗所作这一流行看法提出异议。陈炳藻的研究成果引起红学界较大的反响,为红学乃至其他文学著作的研究开辟了一个新的研究途径。从其他学科对统计学的推动作用看,不少现代统计方法来源于一些实质性学科的定量研究活动。例如,最小二乘法与正态分布理论源于天文观察误差分析,相关与回归分析源于生物学研究,主成分分析与因子分析源于教育学与心理学的研究,抽样调查方法源于政府统计调查资料的收集等。随着科学的进步,在定性分析的基础上进一步开展定量分析已成为现代科学的普遍发展趋势,各学科会越来越需要统计学提供相应的方法论支持。

其次,国际统计学界的主流也从原来的偏重数理统计学的研究向更加重视应用统计研究

转变。目前国际统计学会共有 7 个分会:贝努里数理统计与概率学会、国际统计计算协会、国际官方统计协会、国际调查统计学家协会、国际统计教育协会、国际工商统计学会、欧文·费雪货币与金融统计委员会(或称中央银行统计的欧文·费雪委员会)。这 7 个分会中,数理统计与概率学会有着最为悠久的历史,其他几个分会成立得较晚,如官方统计协会、工商统计学会、货币与金融统计委员会都是在 20 世纪 80 年代以后才陆续建立的。这些分会的研究领域涉及面很广,其中相当一部分属于社会经济领域,其所采用的方法也不仅限于概率统计方法。我国著名统计学家陈希孺院士在就“统计学发展前景展望”答记者问时曾提到:“统计学研究应努力与其他实用学科结合而形成交叉或边缘学科。这一点目前已有一定的表现,如生物统计、医药统计、工业统计、金融统计等,都是当前发展很快的热点。”国际知名的统计学家劳(C. R. Rao)也提出:“统计学基本上是寄生的:靠研究其他领域内的工作而生存。”主张统计学必须与另一门专门的学问相结合,才有可能做出有重要意义的成果。国际知名的统计学家休伯(Huber)甚至认为,统一的统计学将会因为与其他学科结合发展而分裂成许多并行的学科,好比一个大国分裂成一些小国,并把这称为统计学的巴尔干化。对这种趋势,不少统计学家持肯定态度,并不认为这将导致一般统计学的消亡。统计学与各实质性学科更紧密地结合,并根据各自领域的特点和问题,开发和研究新的统计方法,将推动整个统计学科的发展。一般应寓于特殊应用领域之中,那种抽象地不与实际领域相结合去研究统计的路将越走越窄。

再次,统计学与计算机科学和信息科学的结合越来越紧密。在统计数据的采集方面,电话和互联网的迅速普及以及计算机技术的进步,为统计调查提供了电脑辅助的电话调查和网络调查等新的统计调查方式,不仅使统计调查的效率大大提高,还可能带来统计数据收集方式的重大变革。在统计计算方面,统计科学与计算机技术结合,产生了不少新的统计和计量分析的软件和编程软件。如用于编程的 R 语言、用于统计分析的 SAS 软件和 SPSS 软件、用于计量分析的 EVIEWS 软件等。近年来,针对现实生活中海量数据的出现,统计科学与计算及科学和数学等相结合还产生了数据挖掘(如何从海量数据中获取有用信息)等新的研究领域。

纵观统计学发展的历史,我们可以发现:统计学最初是从设置指标研究社会经济现象的数量开始的。随着社会的发展,为了适应实践的需要,统计方法和理论不断丰富和完善,统计学也在不断发展和演变。从当前世界各国的状况来看,统计学已经成为研究客观世界数量规律的最有力工具,它既研究确定现象的数量方面,也研究随机现象的数量方面。统计学的作用与功能从描述事物现状、反映事物规律,向进行抽样推断、预测未来、提供决策依据的方向扩张。统计学自身也从单一的实质性社会科学演变成横跨社会科学领域和自然科学领域的多科性的方法论科学。

三、理论统计学和应用统计学

如前所述,“统计学是研究如何测定、收集、整理、归纳和分析反映客观现象总体数量的数据,以帮助人们正确认识客观世界数量规律的方法论科学”。从横向看,各种统计学都具有上述共同点,因而能够形成一个学科“家族”。从纵向看,统计学方法应用于各种实质性科学,同它们相结合,产生了一系列专门领域的统计学(参见图 1-2)。

由此可见,现代统计学可以分为两大类:一类是以抽象的数量为研究对象,研究一般的收集数据、整理数据和分析数据方法的理论统计学;另一类是以各个不同领域的具体数量为研究对象的应用统计学。

理论统计学把研究对象一般化、抽象化,以数学中的概率论为基础,从纯理论的角度对统