

经济与管理类统计学系列教材

计量经济学

庞皓 主编

科学出版社

北京

内 容 简 介

本书是经济、管理类本科各专业核心课程的教材。本书从我国经济、管理类各专业教学的实际出发,充分借鉴了国内外教材的优点,精选了教学内容,坚持“重思想、重方法、重应用”的原则,避免了烦琐的数学推导和证明,系统介绍了计量经济学的基本理论、基本思想、基本方法及其应用,包含了教育部经济学学科教学指导委员会制定的经济学科本科计量经济学课程基本要求的全部内容。本书特别突出计量经济学的实际应用,每一章都有实际的经济案例,与普遍应用的 EViews 软件紧密结合,并且专门讨论了应用计量经济学方法作实证项目研究的一般方式。

本书适用于经济、管理类各专业教师和本科学生,也特别适合于自学计量经济学的读者阅读。

图书在版编目(CIP)数据

计量经济学 / 庞皓主编. —北京: 科学出版社, 2007

普通高等教育“十一五”国家级规划教材

ISBN 978-7-03-018368-2

I. 计… II. 庞… III. 计量经济学-高等学校-教材 IV. F224.0

中国版本图书馆 CIP 数据核字 (2006) 第 158234 号

责任编辑: 卢秀娟 林 建 / 责任校对: 包志虹

责任印制: 安春生 / 封面设计: 耕者设计工作室

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

印刷

科学出版社发行 各地新华书店经销

*

2007 年 1 月第 一 版 开本: B5 (720×1000)

2007 年 1 月第一次印刷 印张: 24 3/4

印数: 1—19 000 字数: 459 000

定价: 29.00 元

(如有印装质量问题, 我社负责调换〈路通〉)



前言

本书是为高等学校经济管理类各专业本科计量经济学课程编写的教材。中国高等学校开设计量经济学课程已有 20 多年的历史，起初只是部分学校的少数专业开设，1998 年经教育部高等学校经济学学科教学指导委员会讨论决定，把计量经济学纳入了经济学类所有专业必修的核心课程，此后计量经济学更加受到经济学类各专业的普遍重视。在全国各高校中，不仅经济学类各专业已普遍开设了计量经济学课程，而且一些管理类专业也十分重视这门课程的学习。经过 20 多年的努力，中国高等学校的计量经济学教学已经有了长足的进步。目前，不仅引进或翻译了许多国外的计量经济学教材，而且国内也编写了不少教材，与 20 年前刚开设计量经济学课程时教材奇缺的状况相比，已经有了很大改善。但是，从中国高等学校经济管理类各专业学生的实际出发，作为各专业的共同基础课，应该怎样合理地组织教学内容，怎样用有限的课时使学生既掌握计量经济学的基本理论和方法，又具备运用计量经济学知识分析实际经济问题的能力，还需要认真地加以研究。现在编写计量经济学教材，已不是解决教材的有无问题，而是要在总结多年教学经验的基础上，努力提高教材的质量，编写出最适合于经济管理各专业本科教学使用的教材。

本书充分借鉴了国内外教材的优点，总结了编者多年来在财经院校从事计量经济学教学的经验和体会，是在对过去多次编写的教材反复思考、多方提炼的基础上，重新编写而成的。目标是力图做到“教师最好教，学生最好学”。与其他同类教材相比，本书有一些明显的特点：

- 1) 从经济管理类各专业的实际出发，精选了教学内容。本科阶段的计量经济学课程的目标，应当定位在使学生掌握计量经济研究的最基本方法，并能够运用这些方法解决实际的经济问题。大学本科的计量经济学课程一般都只安排一学

期的入门课程，只能以经典计量经济学的内容为主，适当概要性地介绍一些新发展的方向。经典计量经济学应用最为普遍，也是更高层次计量经济学课程的重要基础，符合财经院校绝大多数本科专业教学的实际要求，非经典计量经济学的内容应该放到更高层次的教材中去。本书中章节标题未用脚注注明的部分，是本科计量经济学教学的最基本要求。考虑到全国各学校、各专业的教学要求有一定差异，本书也安排了部分选讲内容，在相应的章节标题中以脚注标出，供本科教学选择使用，但跳过这些内容，并不影响对计量经济学基础内容的系统学习。

2) 坚持“重思想、重方法、重应用”的原则，特别注重基本思想、经济背景、基本方法和实际应用。计量经济学是一门经济学课程，并不是数学课。多年来，学生反映计量经济学课程较难，教材看不懂，其原因是教学内容和教材的写法过于数学化。本书尽可能避免烦琐的数学推导，少数必要的数学推导和证明也是放到附录中，供选择阅读，使之更加适应更多经济管理专业学生的要求。

3) 为教学创造良好的条件和环境，根据我们的教学体会，在每一章的开始都设置了从实际经济背景出发提出的“引子”，目的是从实际应用的角度提出本章将要讨论的主要问题，而不是从概念到概念，抽象地讲理论和方法。通过一些案例分析说明相应章节讨论的主要内容如何通过 EViews 计算机软件去实际运用。计量经济学中概念和公式较多，为有利于教师对本章的学习内容作总结和学生作复习，每一章的最后除了思考题和练习题以外，还提供了小结，多数小结中以表格形式列出了各自章节的主要公式。

4) 本书与普遍应用的 EViews 计算机软件紧密结合，书中讲述的所有方法都要求在 EViews 软件上实现。改变了过去单独介绍软件的做法，将 EViews 软件的学习与各章案例分析有机结合，使学生在实际运用中学习 EViews 的操作方法。

5) 许多学生反映学习了计量经济学后不知该怎么运用，对计算的结果难以做出合理的解释。为了培养学生应用计量经济学方法独立解决实际经济问题的能力和素质，本书改变了其他教材介绍若干宏观经济应用模型的做法，在第十二章专门讨论应用计量经济学方法做实际项目研究的一般方式，指导学生通过完成“课程论文”，去自己体验计量经济学方法的实际应用，并提高计量经济分析的实际应用能力。从 2000 年起，我们就在计量经济学本科的教学中全面采用了这种教学方式，取得了较好的效果。在计量经济学教材中做这样的改革，是在总结教学实践经验基础上的一种探索。

本书第一至三章由西南财经大学教授庞皓编写，第四章由中南财经大学教授徐映梅博士编写，第五、十一章由西南财经大学教授李南成博士编写，第六章由中南财经大学副教授李占风编写，第七、十章由西南财经大学教授史代敏博士编写，第八、九、十二章由西南财经大学教授黎实博士编写。庞皓教授对全书做了

修改。本书主审、山西财经大学教授杭斌认真审阅了全书，提出了许多很好的修改意见。

本书适合作为高等院校经济管理类本科各专业计量经济学课程的教材，如果适当考虑供教学选择的内容，也可作为非数量经济、非统计专业研究生的辅助教材。同时，本书还特别适合自学计量经济学的读者学习。

由于编者水平有限，书中错误在所难免，恳请广大读者批评指正。

编 者

2006年1月

目录

前 言

第一章

导论.....	1
第一节 什么是计量经济学	1
第二节 计量经济学的研究步骤	5
第三节 变量、参数、数据与模型	10
本章小结	14
思考题	14

第二章

简单线性回归模型.....	16
第一节 回归分析与回归函数	17
第二节 简单线性回归模型参数的估计	29
第三节 拟合优度的度量	39
第四节 回归系数的区间估计和假设检验	42
第五节 回归模型预测	48
第六节 案例分析	54
本章小结	61

思考题	63
练习题	64
附录 2.1 简单线性回归最小二乘估计最小方差性质的证明	68
附录 2.2 σ^2 最小二乘估计的证明	70

第三章

多元线性回归模型	72
第一节 多元线性回归模型及古典假定	74
第二节 多元线性回归模型的估计	78
第三节 多元线性回归模型的检验	84
第四节 多元线性回归模型的预测	90
第五节 案例分析	92
本章小结	96
思考题	98
练习题	98
附录 3.1 多元线性回归最小二乘估计无偏性的证明	101
附录 3.2 多元线性回归最小二乘估计最小方差性的证明	101
附录 3.3 残差平方和 $\sum e_i^2$ 的均值为 $(n-k)\sigma^2$ 的证明	103

第四章

多重共线性	104
第一节 什么是多重共线性	106
第二节 多重共线性产生的后果	108
第三节 多重共线性的检验	112
第四节 多重共线性的补救措施	115
第五节 案例分析	119
本章小结	123
思考题	124
练习题	125

第五章

	异方差性	129
第一节	异方差性的概念	130
第二节	异方差性的后果	132
第三节	异方差性的检验	133
第四节	异方差性的补救措施	138
第五节	案例分析	141
	本章小结	148
	思考题	149
	练习题	149
附录 5.1	在异方差性条件下参数估计统计性质的证明	154
附录 5.2	对数变换后残差为相对误差的证明	156

第六章

	自相关	157
第一节	什么是自相关	158
第二节	自相关的后果	160
第三节	自相关的检验	164
第四节	自相关的补救	168
第五节	案例分析	171
	本章小结	175
	思考题	176
	练习题	176
附录 6.1	存在自相关时参数估计值方差的证明	180

第七章

	分布滞后模型与自回归模型	182
第一节	滞后效应与滞后变量模型	183
第二节	分布滞后模型的估计	185
第三节	自回归模型的构建	191

第四节	自回归模型的估计	196
第五节	案例分析	199
	本章小结	208
	思考题	210
	练习题	210

第八章

	虚拟变量回归	215
第一节	虚拟变量	216
第二节	虚拟解释变量的回归	218
第三节	虚拟被解释变量	227
第四节	案例分析	234
	本章小结	237
	思考题	239
	练习题	239

第九章

	设定误差与测量误差	243
第一节	设定误差	244
第二节	设定误差的检验	249
第三节	测量误差	252
第四节	案例分析	255
	本章小结	259
	思考题	260
	练习题	260
附录 9.1	$\hat{\alpha}_2$ 概率极限性质的证明	263
附录 9.2	参数 $\hat{\alpha}_2$ 一致性的证明	263
附录 9.3	有测量误差模型参数估计结果的推导	264

第十章

	时间序列计量经济模型	265
第一节	时间序列计量经济分析的基本概念	266
第二节	时间序列平稳性的单位根检验	268
第三节	协整	273
第四节	案例分析	277
	本章小结	283
	思考题	284
	练习题	284

第十一章

	联立方程组模型	288
第一节	联立方程模型及其偏倚	289
第二节	联立方程模型的识别	296
第三节	联立方程模型的估计	307
第四节	案例分析	311
	本章小结	317
	思考题	319
	练习题	319
	附录 11.1 联立方程偏倚的证明	323

第十二章

	实证项目的计量经济研究——课程论文分析	325
第一节	实证项目研究的选题	326
第二节	模型设定与数据处理	330
第三节	计量经济分析	336
	附录 12.1 实证项目研究（课程论文）示例	340
	主要参考文献	352

附录 统计用表.....	353
表 1 标准化正态分布下的面积	354
表 2 t 分布的百分点	355
表 3 F 分布的上端百分点	356
表 4 χ^2 分布的上端百分点	364
表 5 (a) 德宾-沃森 d 统计量 (在 0.05 显著性水平上 d_L 和 d_U 的显著点)	366
表 5 (b) 德宾-沃森 d 统计量 (在 0.01 显著性水平上 d_L 和 d_U 的显著点)	371

第一章

导 论

引子

“第二次世界大战后的经济学是计量经济学的时代。”

——P·萨缪尔森 (P·Samuelson)

“在大多数大学和学院中，计量经济学的讲授已成为经济学课程表中最有权威的一部分。”

——R·克莱因 (R·Klein)

第一节 什么是计量经济学

计量经济学是现代经济学的重要分支。为了深入学习计量经济学的理论与方法，有必要首先从整体上对计量经济学做出一些概略性的认识，了解计量经济学的性质、沿革、研究方法以及若干常用的基本概念。

一、计量经济学的产生与发展

计量经济学 (econometrics) 这个词是挪威经济学家、第一届诺贝尔经济学奖获得者弗瑞希 (R. Frisch) 在其 1926 年发表的《论纯经济问题》一文中，按照“生物计量学” (biometrics) 一词的结构仿造出来的。econometrics 一词的本意是指“经济度量”，研究对经济现象和经济关系的计量方法，因此 econometrics 有时也译为“经济计量学”。将 econometrics 译为计量经济学，是为了强调计量经济学是一门经济学科，不仅要研究经济现象的计量方法，而且要研究经济现象发展变化的数量规律。

计量经济学的产生源于对经济问题的定量研究，这是社会经济发展到一定阶段的客观需要。经济现象本来就充满着数量关系，人们很早就探索用定量的方式研究经济问题。早在17世纪，英国经济学家、统计学家威廉·配第在《政治算术》中就运用统计方法研究社会经济问题，主张用“数字、重量和尺度”来阐明经济现象。以后的相当一段时间，经济学家们也力图运用数学方法研究经济活动，用数学语言和公式去表达经济范畴和经济规律。但这都还没有形成计量经济学。计量经济学作为经济学的一门独立学科被正式确立，其标志一般认为是1930年12月R. 弗瑞希和丁伯根（J. Tinbergen）等经济学家在美国克里富兰成立国际计量经济学会。

第二次世界大战以后，计量经济学在西方各国的影响迅速扩大，发展成为经济学的重要分支。特别是20世纪40~60年代，经典计量经济学逐步完善并得到广泛应用。美国著名经济学家、诺贝尔经济学奖获得者萨缪尔森认为：“第二次世界大战后的经济学是计量经济学的时代”。事实上，在诺贝尔经济学奖获得者中，相当一部分都是计量经济学家。

20世纪70年代以来，计量经济学的理论和应用又进入一个新的阶段。首先是计算机的广泛应用和新的计算方法大量提出，计量经济模型的规模越来越大。更重要的是，非经典计量经济学的理论和应用有了新的突破。微观计量经济学、非参数计量经济学、时间序列计量经济学和动态计量经济学等的提出，使计量经济学产生了新的理论体系，协整理论、面板数据、对策论、贝叶斯方法等理论在计量经济学中的应用已成为新的研究课题。

应该看到，计量经济学的发展是与现代科学技术成就结合在一起的，它反映了社会化大生产对各种经济因素和经济活动进行数量分析的客观要求。经济学从定性研究向定量分析的发展，是经济学逐步向更加精密、更加科学发展的表现。正如马克思强调的：一种科学只有成功地运用了数学以后，才算达到了完善的地步。因此，诺贝尔经济学奖获得者、经济学家克莱因认为：“计量经济学已经在经济学科中居于最重要的地位。”

计量经济学的一个重要特点是它自身并没有固定的经济理论，计量经济学中的各种计量方法和技术，大多来自数学和统计学。我们只要坚持以科学的经济理论为指导，紧密结合中国经济的实际，就能够使计量经济学的理论与方法在中国的经济理论研究和现代化建设中发挥重要的作用。

二、计量经济学的性质

计量经济学的奠基人弗瑞希指出：计量经济学是“统计学、经济学和数学的结合”，“三者结合起来，就有力量，这种结合便构成了计量经济学”。

《美国现代经济词典》认为：计量经济学是用数学语言来表达经济理论，以

便通过统计方法来论述这些理论的一门经济学分支。

萨缪尔森、库普曼斯、斯通等著名经济学家在 1954 年计量经济学家评审委员会的报告中认为：“计量经济学可定义为，根据理论和观测的事实，运用合适的推理方法使之联系起来同时推导，对实际经济现象进行数量分析。”

尽管这些经济学家对计量经济学定义的表述各不相同，但可以看出，计量经济学不是对经济的一般度量，它与经济理论、统计学、数学都有密切的关系。事实上，计量经济学是以经济理论和经济数据的事实为依据，运用数学、统计学的方法，通过建立数学模型来研究经济数量关系和规律的一门经济学科。应当注意，计量经济学所研究的主体是经济现象及其发展变化的规律，所以它是一门经济学科。计量经济学当然会运用大量的数学方法，特别是许多数理统计方法，但数学在这里只是工具，而不是研究的主体。

计量经济学的目的是要把实际经验的内容纳入经济理论，确定表现各种经济关系的经济参数，从而验证经济理论，预测经济发展的趋势，为制定经济政策提供依据。为此，计量经济学不仅要寻求经济计量分析的方法，而且要对实际经济问题加以研究，要解决达到上述目的的理论和方法论问题。这样，计量经济学分成了两种类型：即理论计量经济学和应用计量经济学。

理论计量经济学研究如何建立合适的方法去测定由计量经济模型所确定的经济关系。现实的经济活动和经济关系异常复杂，一般来说，各种经济变量之间并不是精确的函数关系，经济变量间的数量关系不是那么确定，也就是说，模型中往往包含一些随机的无法直接控制的因素，所以理论计量经济学要较多地依赖数理统计学方法。除了介绍计量经济模型普遍应用的参数估计方法与检验方法以外，由于经济现象的复杂性，各种实际的经济关系不一定都服从一般的统计规律，理论计量经济学还需研究当一般的统计假定条件不完全满足时将会产生的结果，并寻求解决这些问题的专门方法，也就是说还会面临许多特殊的经济问题，形成一些专门的计量经济方法。所以理论计量经济学是适合于经济关系计量的方法论学科。

应用计量经济学是运用理论计量经济学提供的工具，研究经济学中某些特定领域的经济数量问题，如生产函数、消费函数、投资函数、供给函数、劳动就业等。应用计量经济学以建立应用计量经济学模型为主要内容，强调应用模型的经济学和经济统计学基础，侧重于建立与应用模型过程中实际问题的处理。应用计量经济学研究的是具体的经济现象和经济关系，研究它们在数量上的联系及其变动规律性。除了计量经济方法以外，应用计量经济学更多地要依据经济学理论所确定的经济规律，而且要依据经济统计提供的反映现实经济现象和经济关系的观测数据，运用计量经济模型分析经济结构，预测经济的发展趋势，对经济政策做定量的评价。

三、计量经济学与其他学科的关系

从前面的讨论可以看出，计量经济学是与经济学、经济统计学及数理统计学都有关系的交叉学科。但计量经济学又不是这些学科的简单结合，它与这些学科既有联系又有区别。

计量经济学研究的主体是经济现象和经济关系的数量规律，这决定了计量经济学应当以经济学提供的理论原则和揭示的经济规律为依据。经济学理论所说明的经济规律，是计量经济学分析经济数量关系的理论依据。离开了经济理论的指导，计量经济学就可能无的放矢，计量经济学的应用也可能会步入歧途。

但是计量经济学并不是盲目地重复经济理论，计量经济学研究是把经济理论与客观现实联系起来分析，计量经济分析的成果或者是对经济理论确定的原则加以验证与充实，或者可以否定某些经济理论原则而做出补充或修改。计量经济学与经济学的明显区别，在于一般的理论经济学主要根据逻辑推理得出结论，主要用文字或符号说明经济现象和过程的本质与规律，大多具有定性的性质。理论经济学有时也会涉及经济现象的数量关系，如说明价格与商品需求量及供应量成正比或反比的关系，但经济理论并不提供这类经济关系数量上的度量，并不说明价格的变动将会使供应量和需求量具体增加或降低多少。计量经济学则要对经济理论所确定的经济关系做出定量的估计，也就是对经济理论提供经验的内容。

经济统计学也研究对经济现象的计量，只不过是侧重于对社会经济现象的描述。经济统计提供的数据，是计量经济学据以估计参数、验证理论的基本依据。离开了经济统计，任何对实际经济问题的经济计量分析都会寸步难行。计量经济学对经济统计的这种依赖性是由经济活动的特殊性决定的。经济现象是人所从事的社会性活动，它不可能像对自然现象的物理实验和化学实验那样，可以在实验室中严格控制其他条件不变，去反复观测某种因素变动对所研究现象的影响。经济现象不可能人为地控制“其他条件不变”，能够做的只是被动地观测客观经济活动的既成事实，也就是分析对实际经济现象观测所得的统计数据。

计量经济学所研究的经济现象并不都呈现为精确的函数关系，计量经济模型中包含了随机误差项，这样模型中的一些变量和参数的估计量都成为了随机变量。数理统计学是研究随机变量统计规律性的学科，所以数理统计学中的回归分析、参数估计、假设检验、方差分析等方法在计量经济学中得到了全面运用，可以说数理统计学是计量经济学的方法论基础。然而，数理统计学只是抽象地研究一般随机变量的统计规律，主要讨论在一定假设条件下一般随机变量的概率分布性质以及特征值的估计与推断。而计量经济学是从具体的经济模型出发，其参数都具有特定的经济意义，研究对模型参数的估计与推断时，不仅要看在数学原理上是否通得过，还要看与实际的经济内容是否一致。而且，在实际经济问题的计量

中，数理统计中一些标准的假定经常不能满足，还需要建立许多专门的经济计量方法。所以，计量经济学并不只是对数理统计方法的简单应用。

作为对计量经济学与其他相关学科关系的总结，可以引述早在 1933 年弗瑞希为《计量经济学》杂志写的发刊词中的一段话：“对经济的数量研究可以从好几个方面着手，但其中任何一个方面就其本身来说都不应该与计量经济学混为一谈。因此，计量经济学与经济统计学决非一码事；它也不同于我们所说的一般经济理论，尽管经济理论大部分都具有一定的数量特征；计量经济学也不应视为数学应用于经济学的同义语。经验表明，统计学、经济理论和数学这三者对于实际理解现代经济生活中的数量关系来说，都是必要的。但任何一种观点本身都不是充分条件，三者结合起来才是强有力的，正是这种结合才构成了计量经济学。”

第二节 计量经济学的研究步骤

运用计量经济学研究经济问题，一般可分为四个步骤：确定变量和数学关系式——模型设定；分析变量间具体的数量关系——估计参数；检验所得结论的可靠性——模型检验；做经济分析和经济预测——模型应用。

一、模型设定

所谓经济模型是指对经济现象或过程的一种数学模拟。社会经济现象和过程是非常复杂的，影响因素众多，经济模型只能把所研究的主要经济因素（表现为经济变量）之间的关系，用适当的数学关系式近似地、简化地表达出来。例如，为了研究居民的消费行为，根据经济学中关于消费行为的理论，认为居民消费支出与其收入成正比例，可将二者的关系表示为以下消费函数

$$Y = \alpha + \beta X \quad (1.1)$$

其中， Y 为居民消费支出； X 为居民家庭收入； α 和 β 为参数。

(1.1) 式中的 β 实际是经济学中的边际消费倾向 (MPC)， β 作为斜率系数是消费增加量 ΔY 与收入增加量 ΔX 的比例，即 $\beta = \Delta Y / \Delta X$ 。然而，在现实的经济生活中，居民消费支出并不像 (1.1) 式那样，是家庭收入的精确函数。由于还有许多其他未加入模型的因素也会影响居民的消费行为，相同收入的家庭，其消费支出不一定完全相同，所以 (1.1) 式那样的模型还不是适于对实际经济活动做计量分析的计量经济模型。为了把实际居民消费与实际收入水平的关系表现出来，还需要在模型中引入一个随机误差项，即

$$Y = \alpha + \beta X + u \quad (1.2)$$

其中， u 是随机误差项，也称随机扰动项。像 (1.2) 式那样，包含了经济变量、

待确定的参数 α 和 β ，并包含了随机误差项 u 的方程式，才是适于对实际经济活动做计量分析的计量经济模型。计量经济模型可以如 (1.2) 式那样只是一个方程式，这称为单一方程模型。有时，需要用相互联系的若干个方程构成的方程组去描述更为复杂的经济关系，这种计量模型称为联立方程模型。

显然，在建立计量经济模型时，为了简化和计量的方便，通常不可能把所有的因素都列入模型，而只能抓住主要影响因素和主要特征，而不得不舍弃某些因素。同时，模型中变量之间的关系可能设计为线性关系，也可能设计为其他非线性关系。建立模型时，模型中变量的取舍及相互关系形式的设计，一定程度上是决定于研究者的主观认识，当不同的研究者对所研究经济问题的认识有差异时，所使用的模型可能会不完全相同。所以，模型的具体形式是需要研究者去设定的问题。设定模型是计量经济研究的关键步骤，设定计量经济模型既是一门科学，又是一门艺术。建立一个好的计量模型，要靠丰富的专业知识，要有适当的方法，更要靠对建模实践的不断总结。

一般说来，设定一个合理的计量经济模型，主要应注意以下 3 方面的问题：

1) 要有科学的理论依据。建立经济模型是为了反映实际经济活动的规律性，必须对所研究的经济现象的相互关系做科学的理论分析，尽可能使模型真实地反映经济现象实际的依存关系。对别人成功应用过的计量经济模型，也要从经济机理上具体分析，注意模型的应用条件是否符合所研究问题的实际，不应简单地生搬硬套。

2) 模型要选择适当的数学形式。模型的数学形式可以是单一方程，也可以是联立方程。每一个方程可以表现为线性形式，也可以表现为非线性形式。这要根据研究的目的、所研究经济问题的复杂程度以及所掌握的数据资料来决定。可以利用经济学和数理经济学的成果，或利用样本数据绘制变量之间关系的图形作参考。在实际建立模型的过程中，应根据所研究现象相互关系的性质，通过对实际统计资料的试验和分析，经过反复比较，选择尽可能合理的模型数学形式。另外要注意所构造的方程必须是有解的，特别是在建立联立方程模型时，要使内生变量的数目与方程个数相适应。在选择模型数学形式时还应注意，在能够达到研究目的的前提下，应当尽可能选择更为简捷的数学形式，不应只是盲目追求数学形式上的“完美性”。

3) 方程中的变量要具有可观测性。因为只有可观测的变量才可能取得实际的统计数据，也才可能对模型中的参数做出具体的估计。

二、估计参数

参数与变量不同，它是计量经济模型中表现经济变量相互依存程度的那些因

素，通常参数在模型中是一些相对稳定的量。计量经济模型中的参数决定着变量之间的数量关系，一旦参数确定了，整个经济系统的基本结构就确定了。例如，(1.2) 式中作为边际消费倾向的参数 β ，决定着收入与消费的基本结构关系，这种反映经济结构特性的参数也称为结构参数。

在经济总体中，反映经济结构的参数与变量不同，一般来说，参数不能直接观测而且是未知的。我们能够获得的，往往只是所研究总体中变量的若干样本观测数据。由于随机误差项的存在，变量之间的数量关系并不呈现为确定的函数关系，通常也不可能精确地去计算参数的数值。如何通过变量的样本观测数据正确地估计总体模型的参数，这是计量经济学的核心内容。

经过实际样本信息估计出的参数数值称为参数的估计值，但是由于样本毕竟不等于总体，参数的样本估计值并不一定等于总体参数的真实值。如果用一定的方法能够获得对参数估计过程的公式，这种公式则称为参数的估计式或估计量。参数估计式是模型中变量样本观测值的代数式，只要将变量的样本观测值直接代入估计式，即可得到参数的估计值。如何去确定满足计量经济要求的参数估计式，是理论计量经济学的主要内容之一。

三、模型检验

模型中的参数被估计以后，一般说来这样的模型还不能直接加以应用，还需要对估计的计量经济模型做某些检验，其原因是多方面的。首先，在设定模型时，对所研究经济现象规律性的认识可能并不充分，所依据的经济理论对所研究对象也许还不能做出正确的解释和说明。或者虽然经济理论是正确的，但可能我们对问题的认识只是从某些局部出发，或者只是考察了某些特殊的样本，以局部去说明全局的变化规律，可能导致偏差。其次，我们用以估计参数的统计数据或其他信息可能并不十分可靠，或者较多地采用了经济突变时期的数据，不能真实代表所研究的经济关系，或者由于样本太小，所估计的参数只是抽样的某种偶然结果。此外，我们所建立的模型、采用的方法、所用的统计数据，都有可能违反计量经济的基本假定，这也可能导出错误的结论。所谓模型检验，就是要对模型和所估计的参数加以评判，判定在理论上是否有意义，在统计上是否有足够的可靠性。

对计量经济模型的检验主要应从以下 4 方面进行。

1. 经济意义的检验

模型中的变量和参数都有特定的经济意义，经济理论通常对这些变量以及参数的符号和取值范围做出了理论说明，如果所估计的模型与经济理论完全相符，则说明我们所观测的事实证实了这种理论；如果所估计的模型与经济理论不相符，一般来说应当舍弃所估计的模型，设法从模型设定、估计方法、统计数据等

方面找出导致错误结论的原因。

但是应强调，实践是检验真理的唯一标准。任何经济学理论，只有当它成功地解释了过去，才能为人们所接受。如果经过反复研究，证明计量经济模型和估计的参数完全正确，而是经济理论本身不完备，这时则应提出修正经济理论的建议。所以，计量经济学模型对检验经济理论、发现和发展经济理论也有重要意义。

2. 统计推断检验

模型的参数是用变量的观测值估计的，为了检验参数估计值是否抽样的偶然结果，需要运用数理统计中的统计推断方法，对模型及参数的统计可靠性做出说明。对计量经济模型的统计推断检验，包括对模型的拟合优度的检验、用假设检验和方差分析方法对变量显著性的检验等。

3. 计量经济学检验

计量经济学检验主要是检验模型是否符合计量经济方法的基本假定。例如，检验模型中变量是否存在多重共线性，检验模型中的随机扰动项是否存在自相关和异方差性，检验模型是否存在可识别性，检验模型中经济变量是否存在平稳性等。当模型违反计量经济方法的基本假定时，通常的计量经济方法将失去效用或将导致错误的结论，这时必须对模型做必要的处理，并重新估计模型的参数。

4. 模型预测检验

这是指将估计了参数的模型用于实际经济活动的预测，然后将模型预测的结果与经济运行的实际结果相对比，以此检验模型的有效性。

四、模型应用

经过估计参数和模型检验，确认为可靠的计量经济模型，才可以用于实际的经济计量分析。计量经济模型主要可以用于经济结构分析、经济预测和政策评价等几个方面。

所谓经济结构分析，是指用已经估计出参数的模型，对所研究的经济关系进行定量的考察，以说明经济变量之间的数量比例关系。也就是说，分析当其他条件不变时，模型体系中的解释变量发生一定的变动对被解释变量的影响程度。常用的经济结构分析方法有边际分析、弹性分析、乘数分析、比较静力学分析等。例如，说明一个国家或地区国民总收入 Y 与总消费支出 C 关系的消费函数模型

$$C = \alpha + \beta Y + u \quad (1.3)$$

通常这是一条斜率为正值且小于1的直线。模型中的参数 β 的经济意义是边际消费倾向 $MPC = \Delta C / \Delta Y$ ，假如估计的参数为 $\beta = 0.8$ ，这说明国民总收入每增加1亿元，总消费支出将增加0.8亿元，这是宏观经济结构分析中很有意义的数。在此基础上还可进行边际储蓄趋向和收入增长的乘数分析，因为边际储蓄趋

向 $\Delta S/\Delta Y=1-MPC$ ，收入增长乘数 M 为

$$M = 1/(1 - MPC) \quad (1.4)$$

若已估计出 $\beta = MPC = 0.8$ ，则乘数 $M = 5$ 。这说明当投资增加 1 个单位时，将导致国民总收入增加 5 个单位，这又为经济分析提供了重要的定量信息。

所谓经济预测，是指利用估计了参数的计量经济模型，由已知的或预先测定的解释变量，去预测被解释变量在所观测的样本数据以外的数值。计量经济模型本身就是试图从已经发生的经济活动中找出变化规律，然后把这种规律用于样本以外数据的预测。经济预测可以是对被解释变量未来时期的动态预测，也可以是对被解释变量在不同空间状况的空间预测。

所谓政策评价，是利用计量经济模型对各种可供选择的政策方案的实施后果进行模拟测算，从而对各种政策方案做出评价。在这种情况下，我们是把计量经济模型当作经济运行的“实验室”，去模拟所研究的经济体系，分析整个经济体系对各种假设的政策条件的反映。在实际的政策评价时，经常把模型中的某些变量或参数视为可用政策调整的“政策变量”，然后分析“政策变量”的变动对被解释变量的影响。

显然，计量经济模型应用的经济结构分析、经济预测和政策评价三个方面有

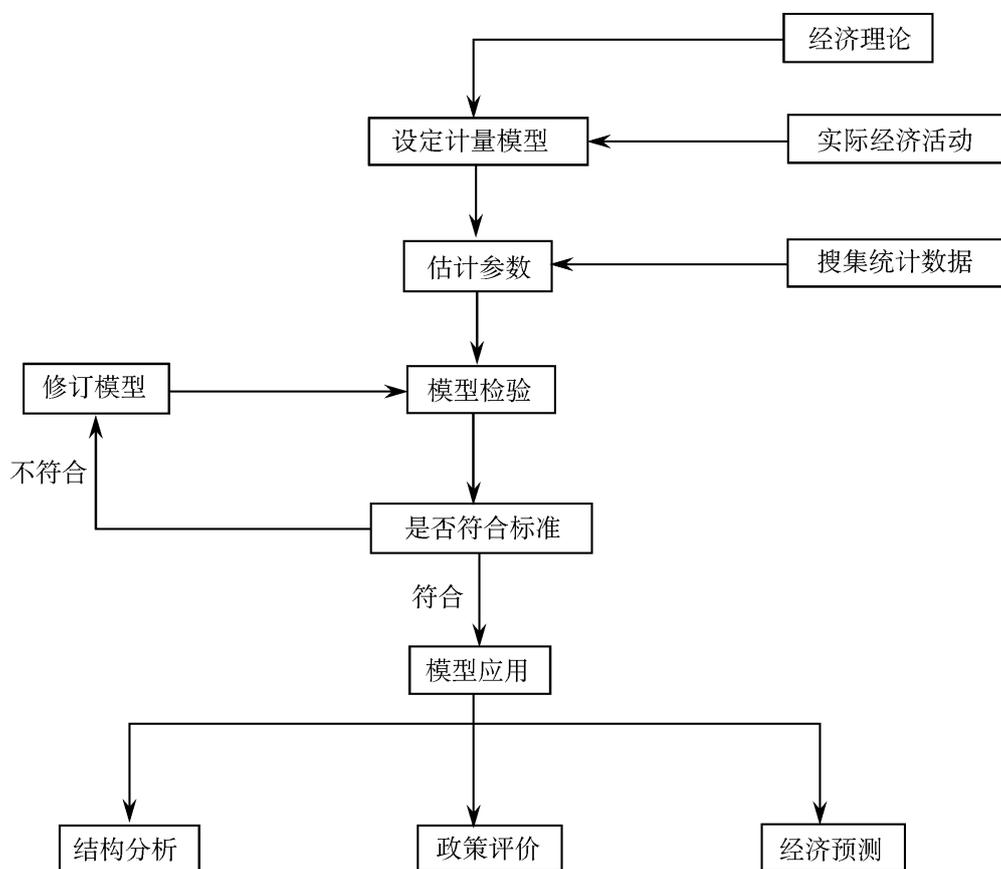


图 1.1 计量经济学的研究过程

密切关系，经济结构分析的结果，可用于经济预测，经济预测的结果是政策评价的依据，而政策评价本身，实际就是一种条件预测。

综上所述，完整的计量经济研究过程可以表示为图 1.1。

第三节 变量、参数、数据与模型

一、计量经济模型中的变量

计量经济模型有多种构成因素，其中一些在不同的时间或空间有不同的状态，会取不同的数值，并且是可以观测的因素，这类因素称为经济变量。例如，(1.2) 式中的居民家庭收入 X 和居民消费支出 Y 都是经济变量。

计量经济模型中的变量可分为若干类型。从所描述的经济活动形态看，经济变量可分为流量和存量。某些变量具有时间维度，是按一定时期测度的总量，它们是一定时期内累计发生的数量，如国内生产总值、投资量、消费量等，这类变量反映的是经济活动的“流量”。另一些变量不具时间维度，是在一定时点上测度的总量，它们表明某一时点存在状态的总量，如金融资产、金融负债等，这类总量反映的是经济“存量”。

从变量的因果关系上，可分为解释变量 (explanatory variable) 和被解释变量 (explained variable)。在模型中，解释变量是变动的的原因，被解释变量是变动的结果。被解释变量是模型要分析研究的对象，也常称为“应变量” (dependent variable)、“回归子” (regressand) 等，如式 (1.2) 中的消费支出 Y 。解释变量也常称为“自变量” (independent variable)、“回归元” (regressor) 等，是说明应变量变动主要原因的变量，如式 (1.2) 中的居民家庭收入 X 。表述被解释变量和解释变量的术语较多，为了表述上尽量一致，避免产生混淆，本书中统一使用“解释变量”表示应变量变动原因的变量；而用“应变量”或“被解释变量”表示分析研究的对象，即作为变动结果的变量。

从变量的性质，又可把变量分为内生变量和外生变量。内生变量是其数值由模型所决定的变量，内生变量是模型求解的结果，如式 (1.2) 中的消费支出 Y 。外生变量是其数值由模型以外决定的变量，如式 (1.2) 模型中的家庭收入。在计量经济模型中，外生变量数值的变化能够影响内生变量的变化，而内生变量却不能反过来影响外生变量。在内生变量中，有一些是过去时期的内生变量或称滞后的内生变量。例如，在研究消费—收入模型时，可能涉及上一期的收入或上一期的消费支出。又如，在研究某地区居民收入对消费的影响时，可能涉及全国居民的收入量。这种过去时期的、滞后的或更大范围的内生变量，不受本模型研究范围的内生变量的影响，但能够影响我们所研究的本期的内生变量，这种内生

变量称为前定内生变量。在模型中前定内生变量的作用视同于外生变量，并与外生变量一起称为前定变量。在单一方程模型中，前定变量一般作为解释变量，内生变量一般作为被解释变量或应变量，而在联立方程模型中内生变量既可作为应变量，又可作为解释变量。

确定模型中的变量，是建立计量经济模型的重要环节，应变量要选择最能反映所研究对象变动情况的变量。例如，最能反映居民消费行为的是“居民消费支出”，最能够反映全社会生产成果总量的是“国内生产总值（GDP）”……解释变量则应尽量选择最能够说明应变量变动的主要原因，并能够独立影响应变量的那些变量，次要的变动原因应被归入到随机扰动项中。对变量的选择还要考虑可观测性，有的因素虽然对应变量有重要影响，但是无法获取其观测值。例如，家庭财产可能对家庭消费有影响，但家庭的财产数量很难取得数据，这类无法观测的因素不能实际度量，也不宜作为变量列入计量经济模型。

二、参数估计的方法

计量经济模型中的参数一般是未知的，需要根据样本信息去加以估计。估计模型中参数的方法有很多种。例如，对于单一方程模型，最常用的是普通最小二乘法、极大似然估计法等。对于联立方程模型常用二段最小二乘法和三段最小二乘法等去估计参数。这些估计方法都是建立在一定假设前提的基础上的，当估计条件不完全满足时，还需要一些特殊的估计方法。由于抽样波动的存在，加之前提条件、估计方法及所确定的估计式不一定那么完备，所得到的参数估计值与总体参数的真实值并不一致，这就要求所得到的参数估计值应符合“尽可能地接近总体参数真实值”的准则。在各种条件下如何寻求模型参数合理的估计方法，是计量经济学研究的主要内容。不过，在理论计量经济学中并不侧重于直接研究参数估计值本身，而是着重于论述所导出的参数估计式是否符合“尽可能地接近总体参数真实值”这样的准则。通常选择参数估计式时应考察其无偏性、最小方差性等统计性质，或者考察大样本时的统计性质。

在实际的计量经济研究中，运用样本数据对参数的估计与检验，将面临很大的计算工作量。现在这方面的工作已经可以由各种有关的计算机应用软件来实现，计算机应用软件已成为学习计量经济学必不可少的部分。本书采用的计算机应用软件是应用十分广泛的 EViews (econometrics views)，将结合每一章的内容介绍 EViews 的使用方法。

三、计量经济学中应用的数据

估计计量经济模型参数的基本依据，是通过对所研究经济变量实际观测所取得的数据。数据是对客观事物信息的一种反映，这种信息如以某种量的标志显现

出来就称其为数据。在计量经济研究中使用的数据，主要是各种经济统计数据，也可以是通过专门调查取得的数据，还可以是人为构造的数据。可用于估计参数的数据主要有以下几类：

1) 时间序列数据 (time series data)。把反映某一总体特征的同一指标的数据，按照一定的时间顺序和时间间隔 (如月度、季度、年度) 排列起来，这样的统计数据称为时间序列数据。例如，逐年的国内生产总值和消费支出、逐月的物价指数等。时间序列数据可以是时期数据，也可以是时点数据。

2) 截面数据 (cross-section data)。同一时间 (时期或时点) 某个指标在不同空间的观测数据，称为截面数据。“不同的空间”可以是指不同的地理区域，也可以是指不同的行业、部门或个人，如同一时间不同家庭的收入和消费支出、某一年各个省 (市) 的国内生产总值等。

3) 面板数据 (panel data)。面板数据指时间序列数据和截面数据相结合的数据，如在居民收支调查中收集的对各个固定调查户在不同时期的调查数据，又如全国各省市不同年份的经济发展状况的统计数据，都是面板数据。

4) 虚拟变量数据 (dummy variables data)。时间序列数据和截面数据都是反映定量事实的数据，这是计量经济分析中用得最多的、最基本的数据。但是还有一些定性的事实，不能直接用一般的数量去计量，如政府政策的变动、自然灾害、政治因素、战争与和平状态……。在计量经济研究中常发现，某些客观存在的定性现象确实对所研究的经济变量有明显的影响，需要把它们引入计量经济模型中，这时常用人为构造的虚拟变量数据去表示这类客观存在的定性现象“非此即彼”的状态。通常以 1 表示某种状态发生，以 0 表示该种状态不发生。这样的虚拟变量数据虽然是人为构造的，但反映了客观存在的定性现象，也可以视为一种数据用作模型参数的估计和检验。

以上各种数据虽然都可用于计量经济模型的估计和检验，但是应注意，由于这些数据的性质各不相同，在具体运用时可能不满足某些假定条件而给计量经济分析带来一些影响。例如时间序列数据若是非平稳的，可能造成“伪回归”；截面数据往往存在异方差；利用面板数据的计量经济模型已成为计量经济学研究的专门问题。

除了模型的正确设定以外，能否取得用于实际计量的适合的样本数据，是计量经济研究成败的关键。计量经济分析中使用的数据主要来自于经济统计，常用的数据可从各种统计年鉴等出版物中取得，特殊的数据则需进行专门的调查才能得到。计量经济研究中使用的数据，要力求真实、可靠、完整，数据的质量直接关系到所估计参数的可靠性。对明显失真的数据，应当予以剔除。在经济结构发生重大变革的时期，其统计数据往往不能反映经济变动的真实趋势和规律，这种“经济突变”时期的数据也不宜直接用于估计模型的参数。此外，我们有时很难

直接找到模型所需要的数据，这时可能还需要对能够得到的数据做重新加工，或者试验寻求与所研究的变量高度相关的代用数据，或者对模型的变量与结构加以调整。

计量经济学中利用的数据是可能获得的统计数据，实际的统计数据可能会有观测误差，也可能数据的数量无法满足估计参数的要求，这些数据还可能不满足参数估计方法的基本假定。这样一来，数据可能引发诸如自由度问题、多重共线性、序列相关、异方差性等一系列问题。如何设法解决由数据引起的问题，也是理论计量经济学和应用计量经济学要专门研究的内容。

四、计量经济模型的建立

在计量经济研究中，模型是对实际经济现象或过程的一种数学模拟，再完美的模型也不可能将所有的因素都纳入其中，模型只不过是对可计量的复杂经济现象的一种简化与抽象。因此模型只能在一定的假设前提下，忽略众多次要因素，而突出若干所关注的主要经济变量，把有关经济变量的相互依存关系表现为方程式。模型的建立主要靠对现实经济问题的深入研究，要遵循科学的理论原则，也要运用适当的方法。某些经济变量的相互关系通常可以用来建立计量经济模型，这些关系主要包括以下几种。

1. 行为关系

行为关系指描述决策者经济行为的某些变量与其他变量的关系。例如，居民消费行为与其收入、物价水平等的关系。利用行为关系建立的模型称为行为方程式。

2. 技术（工艺）关系

这反映由科学技术水平决定的经济变量间的数量关系。例如，说明投入的生产要素与产出的生产成果的技术关系，如著名的柯柏-道格拉斯生产函数，产量 Q 与资本投入量 K 、劳动投入产出量 L 的关系为

$$Q = AK^\alpha L^\beta e^u \quad (1.5)$$

其中， A 、 α 、 β 为参数， u 为随机项。

又如，投入产出模型中的生产量 X_j 与消耗量 x_{ij} 间的关系

$$x_{ij} = a_{ij} X_j \quad (1.6)$$

根据生产技术关系建立的模型称为技术方程式。

3. 制度关系

制度关系指经济现象之间由政府政策和规定的制度所决定的关系。例如，销售税的数量决定于销售额和税率，其中税率是由政府规定的。这样建立的模型称为制度方程式。

4. 定义关系

定义关系指根据定义而表达的恒等式。这类关系是由经济理论或客观存在的经济关系决定的恒等关系，例如

$$\text{国内生产总值} = \text{消费} + \text{投资} + \text{净出口}$$

国民经济中许多平衡关系都可以建立恒等关系，这样的模型称为定义方程式。

以上几种方程式中，最重要、最常用的是行为方程式和技术方程式，这两种方程中都有未知参数需要估计，且每个方程说明了经济结构的某一方面，所以这些方程称为结构方程式。

► 本章小结

1. 计量经济学是以经济理论和经济数据的事实为依据，运用数学、统计学的方法，通过建立数学模型来研究经济数量关系和规律的一门经济学科。计量经济学与理论经济学、经济统计学、数理统计学既有区别又有联系。

2. 计量经济研究分为模型设定、参数估计、模型检验、模型运用等四个步骤。

3. 模型的设定主要是选择变量和确定变量间联系的数学形式。适于对实际经济活动做计量分析的计量经济模型应包含经济变量、待确定的参数和随机误差项。行为方程、技术方程、制度方程和定义方程可作为建立模型时参考。

4. 计量经济模型中的变量分为被解释变量（应变量）和解释变量、内生变量和外生变量。

5. 参数是计量经济模型中表现经济变量相互依存程度的因素，通常具有相对稳定性。参数无法直接观测和计算，只能用适当的方法根据变量的样本观测值去估计。参数估计的方法应符合“尽可能地接近总体参数真实值”的准则。

6. 计量经济研究中应用的数据包括时间序列数据、截面数据、面板数据、虚拟变量数据等。

7. 对模型检验包括经济意义检验、统计推断检验、计量经济学检验和模型预测检验。

8. 计量经济模型主要可应用于经济结构分析、政策评价和经济预测。

► 思考题

1.1 怎样理解产生于西方国家的计量经济学能够在中国的经济理论研究和现代化建设中发挥重要作用？

1.2 理论计量经济学和应用计量经济学的区别和联系是什么？

1.3 怎样理解计量经济学与理论经济学、经济统计学的关系？

- 1.4 在计量经济模型中应变量和解释变量的作用有什么不同?
- 1.5 一个完整的计量经济模型应包括哪些基本要素? 你能举一个例子说明吗?
- 1.6 假如你是中央银行货币政策的研究者, 需要你对增加货币供应量促进经济增长提出建议, 你将考虑哪些因素? 你怎样运用计量经济学的研究方法?
- 1.7 计量经济学模型的主要应用领域有哪些?
- 1.8 如果要根据历史经验预测明年中国的粮食产量, 你认为应当考虑哪些因素? 应当怎样来设定计量经济模型?
- 1.9 参数和变量的区别是什么? 为什么对计量经济模型中的参数通常只能用样本观测值去估计?
- 1.10 你能就时间序列数据、截面数据、面板数据、虚拟变量数据分别举出三个实际例子, 并说明这些数据的来源吗?
- 1.11 为什么对已经估计出参数的模型还要进行检验? 你能举一个例子说明各种检验的必要性吗?
- 1.12 为什么计量经济模型可以用于政策评价? 其前提条件是什么?

第二章

简单线性回归模型

引子

中国旅游业总收入将超过 3000 亿美元吗？

根据国家旅游局统计，2004 年中国旅游业快速增长，全年入境人数 1.08 亿人次，分别比 2003 年和 2002 年增长 18% 和 10%；旅游外汇收入可达 250 亿美元，分别比 2003 年和 2002 年增长 43.7% 和 22.6%；国内旅游人数 9.3 亿人次以上，国内旅游收入 4000 亿元以上，比 2003 年和 2002 年分别增长 5.9% 和 3.1% 以上；出境 2800 万人次，分别比 2003 年和 2002 年增长 38.5% 和 68.7%。目前中国旅游业利用外资的规模达到 500 亿美元，占国内各行业吸收外资总额的 11%。中国公民可以组团前往的旅游目的地国家和地区将达到 63 个。现在，中国人均收入已经跨越 1000 美元关口，按照国际经验，将触发国内社会消费结构的升级。居民消费将由实物消费为主走上实物消费与服务消费并重的轨道。在消费结构升级的推动下，中国旅游业将迎来新一轮的长期增长周期。据《中国旅游业发展“九五”计划和 2010 年远景目标纲要》，到 2010 年，中国旅游入境人数将达 6400 万~7100 万人次，国际旅游外汇收入 380 亿~410 亿美元；国内旅游人数将达到 20 亿~25 亿人次，国内旅游收入 10 000 亿~10 500 亿元人民币。这两项合计总产出将达 13 000 亿~14 000 亿元人民币，旅游总收入占 GDP 的比例将达 8%。旅游业已经成为中国第三产业中最具活力与潜力的新兴产业和国民经济中新的增长点。另据世界旅游及旅行理事会（WTTC）预测，未来 10 年间，中国旅游业将保持年均 10.4% 的增长速度，其中个人旅游消费将以年

均 9.8% 的速度增长, 企业/政府旅游的增长速度将达到 10.9%, 到 2020 年, 中国将成为第一大旅游目的地国和第四大客源输出国。从 2004 中国国际旅游交易会上获悉, 到 2020 年, 中国旅游业总收入将超过 3000 亿美元, 相当于国内生产总值的 8%~11%^①。

推动中国旅游业快速发展的原因是多方面的, 是什么决定性的因素能使中国旅游业总收入到 2020 年达到 3000 亿美元? 旅游业的发展与这种决定性因素的数量关系究竟是什么? 显然, 需要寻求一些方法研究相互联系的经济变量之间的数量关系, 对这类问题的研究应当考虑以下几个方面:

- 1) 确定作为研究对象的经济变量 (如中国旅游业总收入);
- 2) 分析影响研究对象变动的主要因素 (如中国居民收入的增长);
- 3) 分析各种影响因素与所研究经济现象的相互关系 (决定相互联系的数学关系式);
- 4) 确定所研究的经济问题与影响因素间具体的数量关系 (需要特定的方法);
- 5) 分析并检验所得数量结论的可靠性 (需要统计检验);
- 6) 运用数量研究结果做经济分析和预测 (对数量分析的实际应用)。

对经济变量相互关系的计量, 最基本的方法是回归分析。回归分析是计量经济学的主要工具, 也是计量经济学理论和方法的主要内容。只有一个解释变量的线性回归模型是最简单的, 称为简单线性回归模型或一元线性回归模型。本章从最简单的一元线性回归模型入手, 讨论在基本假定满足的条件下, 对经济变量关系计量的基本理论和方法, 这也是以后各章的重要基础。

第一节 回归分析与回归函数

一、相关分析与回归分析

(一) 经济变量间的相互关系

许多社会与经济现象, 除了自身的变动以外, 它们相互之间很可能有一定的依存关系。各种经济变量相互之间的依存关系有两种不同的类型: 一种是确定性的函数关系, 另一类是不确定性的统计关系, 也称为相关关系。

当一个或若干个变量 X 取一定数值时, 某一个变量 Y 有确定的值与之相对

^① 国际金融报, 2004-11-25, 2

应，我们称变量间的这种关系为确定性的函数关系。例如，在销售价格 P 不变的情况下，某种商品销售量 X 与销售额 Y 之间的关系可表示为 $Y = PX$ 。一般情况下，确定性的函数关系可表示为 $Y = f(X)$ 。

当一个或若干个变量 X 取一定值时，与之相对应的另一个变量 Y 的值虽然不确定，但却按某种规律在一定范围内变化，我们称变量之间的这种关系为不确定性的统计关系或相关关系，一般可表示为 $Y = f(X, u)$ ，其中 u 为随机变量。例如，居民的可支配收入 X 与居民的消费支出 Y 之间的关系，通常具有相同收入水平居民的消费支出并不完全相同，这时居民可支配收入 X 与消费支出 Y 会呈现为不确定性的相关关系。居民消费支出 X 之所以与居民可支配收入 Y 不呈现为确定性的函数关系，是因为除了居民可支配收入 X 以外，还存在许多其他的因素也会影响居民消费支出 Y 。

变量之间的相关关系可用坐标图又称散点图去描述。例如，变量 X 和 Y 之间关系的散点图可描述为图 2.1。

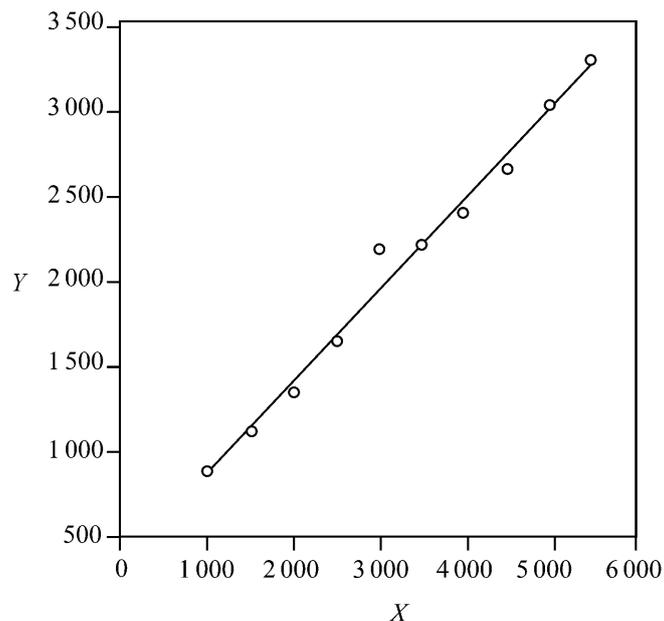


图 2.1 变量相关关系的散点图

由于涉及的变量数量、相关性质及相关程度的不同，变量之间的相关关系可以分为若干类型：

1) 从相关关系涉及的变量数量看。只有两个变量的相关关系，称为简单相关关系，如人的身高与体重之间的相关关系。三个或三个以上变量的相关关系，称为多重相关或复相关，如某种商品的需求量与商品的价格及居民的收入水平之间的相关关系。

2) 从变量相关关系的表现形式看。当变量之间相关关系的散点图中的点接

近一条直线时,称为线性相关;当变量之间相关关系散点图中的点接近于一条曲线时,称为非线性相关。

3) 从变量相关关系变化的方向看。两个变量趋于在同一个方向变化时,即同增或同减,称为变量之间存在正相关。当两个变量趋于在相反方向变化时,即当一个变量增加时,另一个变量却减少,称为变量之间存在负相关。

4) 从变量相关的程度看。当一个变量的变化完全由另一个变量的变化所确定时,称为变量之间完全相关,如前述价格不变时某种商品销售额与销售量之间的关系。在这种情况下,相关关系实际成了函数关系,所以可以把函数关系视为相关关系的特例。

当两个变量的变化相互完全没有关系,即彼此互不影响时,称为二者不相关。两个现象的关系如果介于完全相关和不相关之间时,称为不完全相关,我们研究的相关关系通常都是指的这种不完全的相关关系。

(二) 简单线性相关关系的度量

1. 简单线性相关系数

在各种类型的相关分析中,只有两个变量的线性相关关系的分析是最简单的。两个变量之间线性相关程度可以用简单线性相关系数去度量,这种相关系数是最常用的,也简称为相关系数。对于我们所研究的总体,两个相互联系的变量的相关系数称为总体相关系数,通常用 ρ 表示,总体相关系数 ρ 可用下式计算

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \quad (2.1)$$

其中, $\text{Var}(X)$ 是变量 X 的方差; $\text{Var}(Y)$ 是变量 Y 的方差; $\text{Cov}(X, Y)$ 是变量 X 和 Y 的协方差。

总体相关系数 ρ 反映了总体两个变量 X 和 Y 的线性相关程度,对于特定的总体来说, X 和 Y 的数值是既定的,总体相关系数 ρ 是客观存在的特定数值。然而,当总体较大时,变量 X 和 Y 的全部数值一般不可能去直接观测,所以总体相关系数一般是不能直接计算的未知量。通常可能做到的是从总体中随机抽取一定数量的样本,通过 X 和 Y 的样本观测值 X_i 和 Y_i 去计算样本相关系数,变量 X 和 Y 的样本相关系数通常用 $r_{X, Y}$ 表示,或简记为 r ,可做以下估计

$$r_{X, Y} = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{\sqrt{n \sum X_i^2 - \left[\sum X_i \right]^2} \sqrt{n \sum Y_i^2 - \left[\sum Y_i \right]^2}} \quad (2.2)$$

或

$$r_{X, Y} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (2.3)$$