

普通高等院校经济管理类“十三五”应用型规划教材
【经济管理类专业基础课系列】



计量经济学基础

BASIS OF ECONOMETRICS

主编 付宏 尹康

免费提供
授课用
电子课件



机械工业出版社
China Machine Press

普通高等院校经济管理类“十三五”应用型规划教材
【经济管理类专业基础课系列】

计量经济学基础

BASIS OF ECONOMETRICS

主编 付宏 尹康

参编 刘亚飞 刘习平 宋来胜 黄璨



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

计量经济学基础 / 付宏, 尹康主编. —北京: 机械工业出版社, 2016.8
(普通高等院校经济管理类“十三五”应用型规划教材·经济管理类专业基础课系列)

ISBN 978-7-111-54494-4

I. 计… II. ①付… ②尹… 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 (2016) 第 182514 号

本书作为计量经济学的入门教材,从经管类专业本科教学的实际出发,强调方法应用,注重计量案例讲解。全书共 9 章,包括统计基础知识回顾、经典回归模型、违背经典假设情形的建模、模型的诊断以及两个计量专题——离散数据和面板数据的建模。书中各章都穿插了相应例子,并给出了 EViews 软件的操作结果,便于读者对理论模型的理解与掌握。

本书侧重计量经济学基础知识介绍,适合作为面向本科生的计量经济学课堂教材使用,也特别适合作为经管类专业学生自学计量经济学的参考书。

出版发行:机械工业出版社(北京市西城区百万庄大街 22 号 邮政编码:100037)

责任编辑:董凤凤

责任校对:殷虹

印刷:三河市宏图印务有限公司

版次:2016年8月第1版第1次印刷

开本:185mm×260mm 1/16

印张:9.75

书号:ISBN 978-7-111-54494-4

定价:25.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

客服热线:(010) 88379210 88361066

投稿热线:(010) 88379007

购书热线:(010) 68326294 88379649 68995259

读者信箱:hzjg@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问:北京大成律师事务所 韩光/邹晓东

计量经济学是现代经济学的一个重要分支，是融经济学、统计学、数学和计算机科学于一体的综合、交叉性学科。它以经济理论和经济数据的事实为依据，以统计知识为基础，以数学模型为手段，以计算机软件为分析工具，是研究经济系统中各种经济变量之间的数量关系和规律的一门经济学科。它可以进行结构分析、经济预测、政策效果分析，是一门极具实用价值的经济学科。

作为从事了多年计量经济学教学工作的教师，我们先后尝试使用过多个版本的计量经济学教材，有国内专家编写的，也有国外引进改编的教材，这些教材都各有特色，但在教学活动中，我们也发现这类教材普遍存在一个问题，对于省属财经类高校的本科生而言，这些教材都显得理论色彩过于浓厚，诸多的数学公式使得学生望而生畏。所以，我们一直希望能编写一本属于我们“自己”的计量经济学教材。我们对本书的期待是：能给经管类专业的本科生提供计量经济学的入门指导，内容设置紧凑、合理，难度适宜，注重案例，强调操作。基于上述原则，我们制定了相应的教材编写大纲。

在本书的编写过程中，参考了《计量经济学基础》（古扎拉蒂著，中国人民大学出版社）、《计量经济学》（詹姆斯 H. 斯托克、马克 W. 沃森著，格致出版社）等多本中外教材，在此向有关作者表示衷心感谢。

本书编写工作分配如下：付宏副教授负责全书的统稿工作，尹康博士承担本书的第 1、2、8 章的编写工作，宋来胜博士承担第 3 章的编写工作，刘习平博士承担第 4、5 章的编写工作，黄璨博士承担第 6 章的编写工作，刘亚飞博士负责第 7、9 章的编写工作。

由于时间和水平的限制，书中的疏漏、错谬之处在所难免，恳请读者批评并提出宝贵意见。希望通过我们的共同努力，提高计量经济学的教学水平。

付宏 尹康

2015 年 4 月 16 日

教学目的

本课程的教学目的是让学生理解和掌握计量经济学的基本原理与方法，重点掌握回归分析工具在经济实证研究中的应用，更进一步鼓励学生尝试对一些复杂类型的经济数据进行建模分析。

教学方式和教学手段

本科阶段的计量经济学教学一直是一个具有挑战性的难题。对理论介绍过多，容易让学生对课程性质产生质疑，以为是在上数学课；如果理论讲解过少、过浅，有些结论又难以解释清楚。因此建议在理论教学的基础上，辅助以案例教学，并配以软件操作演示，让教学过程以应用为导向，使学生通过对案例以及软件操作的理解，加深对计量理论的认识，从而达到较好的教学效果。

学时分配

| 教学内容 | 学习要点 | 学时安排 |
|-----------------|---------------------------------|------|
| 第1章 导论 | 1. 了解计量经济学的发展历史 | 2 |
| | 2. 了解计量经济学的基本概念 | |
| | 3. 掌握计量经济学建模的一般步骤 | |
| 第2章 统计基础知识回顾 | 1. 了解概率、随机变量和概率密度等基本概念 | 4 |
| | 2. 了解多元随机变量中联合密度、边缘密度、条件密度等概念 | |
| | 3. 掌握期望、方差、协方差和相关系数的定义及其运算规则 | |
| | 4. 熟悉几类常见的统计分布 | |
| | 5. 掌握参数估计的基本思想及其构造方法 | |
| | 6. 掌握假设检验的基本思想及其检验的一般步骤 | |
| 第3章 一元线性回归模型 | 1. 了解总体回归方程和样本回归方程 | 6 |
| | 2. 掌握古典回归模型的假定 | |
| | 3. 掌握 OLS 的基本原理 | |
| | 4. 了解系数的估计误差和置信区间 | |
| | 5. 掌握回归模型统计显著性检验的意义和方法 | |
| | 6. 掌握回归模型参数估计和统计检验的 EViews 软件实现 | |

(续)

| 教学内容 | 学习要点 | 学时安排 |
|-----------------------|--|------|
| 第 4 章 多元线性回归模型 | 1. 了解多元线性回归模型的基本形式 | 4 |
| | 2. 掌握多元线性回归模型的设定、估计和模型检验 | |
| | 3. 了解多元线性回归模型的预测 | |
| | 4. 掌握运用多元线性回归模型解决现实的经济问题 | |
| 第 5 章 线性回归模型的扩展 | 1. 了解线性回归模型的其他形式 | 4 |
| | 2. 掌握过原点回归和标准化变量回归模型的特点 | |
| | 3. 掌握对数模型、倒数模型和多项式回归模型的线性转换以及模型的经济意义 | |
| | 4. 掌握虚拟解释变量回归模型中加法模型与乘法模型的应用方法 | |
| 第 6 章 违背经典假设的模型 | 1. 了解违背经典假设的几种情况 | 12 |
| | 2. 掌握多重共线性产生的原因、检验和判断方法以及补救措施 | |
| | 3. 掌握异方差产生的原因、如何检验以及修正的方法 | |
| | 4. 掌握自相关产生的原因、如何检验以及修正的方法 | |
| 第 7 章 计量经济模型的设定与诊断 | 1. 了解模型选择的标准 | 4 |
| | 2. 掌握模型拟合不足和过度拟合产生的后果 | |
| | 3. 掌握因变量和解释变量的测量误差对估计结果的影响 | |
| | 4. 了解嵌套模型和非嵌套模型的区别 | |
| | 5. 掌握对非嵌套模型进行选择的方法 | |
| 第 8 章 定性响应回归模型 | 1. 了解定性响应模型的特征及其分类 | 6 |
| | 2. 掌握线性概率模型的估计方法 | |
| | 3. 掌握 Logit 模型和 Probit 模型的估计方法 | |
| | 4. 了解三类定性响应模型的优劣 | |
| 第 9 章 面板数据模型初步 | 1. 了解面板数据的定义 | 6 |
| | 2. 了解混合 OLS 方法 | |
| | 3. 理解不可观测的异质性对混合 OLS 估计量的影响 | |
| | 4. 理解固定效应估计量和随机效应估计量 | |
| | 5. 理解如何通过 Hausman 检验在随机效应估计量和固定效应估计量之间进行选择 | |
| 学时总计 | | 48 |

目录 Contents

前言

教学建议

第1章 导论 /1

- 1.1 计量经济学发展历史 /1
- 1.2 什么是计量经济学 /2
- 1.3 计量经济学研究的一般步骤 /3
 - 1.3.1 建立理论模型 /3
 - 1.3.2 计量模型的设定 /3
 - 1.3.3 数据的收集与整理 /4
 - 1.3.4 模型的参数估计 /5
 - 1.3.5 模型的检验 /5
 - 1.3.6 模型的应用 /5
- 1.4 软件的使用 /6
- 本章小结 /6
- 练习题 /6

第2章 统计基础知识回顾 /7

- 2.1 随机变量与概率分布 /7
 - 2.1.1 概率、样本空间和随机变量 /7
 - 2.1.2 离散型随机变量的概率分布 /8
 - 2.1.3 连续型随机变量的概率分布 /8
- 2.2 二维随机变量 /10
 - 2.2.1 联合分布和边缘分布 /10
 - 2.2.2 条件分布 /10
- 2.3 随机变量的数值特征：期望和方差 /10
 - 2.3.1 期望 /11

2.3.2 方差 /11

2.3.3 协方差与相关系数 /11

2.3.4 独立性 /12

2.4 几类重要的概率分布 /12

2.4.1 正态分布 /12

2.4.2 卡方分布 /14

2.4.3 t 分布 /14

2.4.4 F 分布 /14

2.5 参数估计 /14

2.5.1 统计推断的基本思想 /14

2.5.2 参数估计量及其评价标准 /15

2.5.3 点估计的构造 /16

2.5.4 区间估计 /18

2.6 假设检验 /18

本章小结 /21

练习题 /21

第3章 一元线性回归模型 /23

3.1 回归分析的基本概念 /23

3.1.1 回归的基本含义 /23

3.1.2 回归与相关 /24

3.1.3 回归与因果 /24

3.1.4 总体回归方程 /24

3.1.5 随机扰动项的意义 /26

3.1.6 样本回归方程 /26

3.1.7 对“线性”的解释 /27

3.1.8 回归模型的基本假定 /28

3.2 一元线性回归模型的估计 /28

- 3.2.1 普通最小二乘法 /28
- 3.2.2 模型估计结果的评价 /29
- 3.2.3 参数估计量的统计性质 /31
- 3.2.4 参数的区间估计 /33
- 3.3 模型的检验 /33
 - 3.3.1 回归系数的显著性检验 /33
 - 3.3.2 回归方程的显著性检验 /34
- 3.4 预测 /35
 - 3.4.1 点预测 /35
 - 3.4.2 区间预测: 均值预测 /35
 - 3.4.3 区间预测: 个值预测 /36
- 3.5 案例分析 /36
 - 3.5.1 研究目的与要求 /36
 - 3.5.2 模型设定 /36
 - 3.5.3 估计参数 /37
 - 3.5.4 模型检验 /39
 - 3.5.5 回归预测 /40
- 本章小结 /41
- 练习题 /43
- 第4章 多元线性回归模型 /45
 - 4.1 多元线性回归模型的设定 /45
 - 4.1.1 多元线性回归模型 /45
 - 4.1.2 多元线性回归模型的矩阵形式 /46
 - 4.1.3 多元线性回归模型的基本假定 /46
 - 4.2 多元线性回归模型的估计 /47
 - 4.2.1 参数的普通最小二乘估计 /47
 - 4.2.2 偏回归系数的含义 /48
 - 4.2.3 参数最小二乘估计的最优性质 /49
 - 4.2.4 随机扰动项方差的估计 /49
 - 4.3 多元线性回归模型的检验 /50
 - 4.3.1 拟合优度与调整拟合优度检验 /50
 - 4.3.2 回归系数的显著性检验 (t 检验) /51
 - 4.3.3 回归方程的显著性检验 (F 检验) /51
 - 4.4 多元线性回归模型的预测 /52
 - 4.4.1 均值预测 /52
 - 4.4.2 个值预测 /53
 - 4.5 案例分析 /53
 - 4.5.1 研究的目的要求 /53
 - 4.5.2 模型设定和数据 /53
 - 4.5.3 参数估计 /54
 - 4.5.4 模型检验 /55
- 本章小结 /56
- 练习题 /56
- 第5章 线性回归模型的扩展 /59
 - 5.1 过原点回归 /59
 - 5.2 标准化变量回归 /60
 - 5.3 可线性化的非线性模型 /61
 - 5.3.1 对数模型 /61
 - 5.3.2 倒数模型 /62
 - 5.3.3 多项式回归 /63
 - 5.4 虚拟解释变量回归 /63
 - 5.4.1 加法模型与方差分析 /64
 - 5.4.2 乘法模型: 交互效应与邹至庄检验 /66
 - 5.5 案例分析 /67
- 本章小结 /68
- 练习题 /68
- 第6章 违背经典假设的模型 /70
 - 6.1 多重共线性 /70

- 6.1.1 定义及特质 /70
 - 6.1.2 对模型估计的影响 /72
 - 6.1.3 多重共线性的检验 /74
 - 6.1.4 补救措施 /75
 - 6.1.5 案例分析 /80
 - 6.2 异方差 /82
 - 6.2.1 异方差的性质 /82
 - 6.2.2 出现异方差对 OLS 估计的影响 /85
 - 6.2.3 异方差的检验 /85
 - 6.2.4 异方差补救措施 /87
 - 6.2.5 案例分析 /90
 - 6.3 自相关 /94
 - 6.3.1 自相关的性质 /94
 - 6.3.2 自相关对 OLS 估计的影响 /95
 - 6.3.3 自相关的检验 /96
 - 6.3.4 自相关的修正 /98
 - 6.3.5 案例分析 /100
 - 本章小结 /103
 - 练习题 /104
- 第 7 章 计量经济模型的设定与诊断 /108
- 7.1 模型选择的标准 /108
 - 7.2 模型误设 /109
 - 7.2.1 不足拟合 /109
 - 7.2.2 过度拟合 /110
 - 7.3 测量误差 /111
 - 7.3.1 因变量的测量误差 /111
 - 7.3.2 自变量的测量误差 /112
 - 7.4 嵌套与非嵌套模型 /114
 - 7.4.1 基于信息准则的判别 /114
 - 7.4.2 基于统计检验的判别 /115
 - 本章小结 /115
- 练习题 /116
- 第 8 章 定性响应回归模型 /117
- 8.1 定性响应模型的性质 /117
 - 8.2 线性概率模型 /118
 - 8.2.1 随机扰动项 ϵ_i 的非正态分布 /119
 - 8.2.2 干扰项的异方差 /119
 - 8.2.3 不满足 $0 \leq E(Y_i | X_i) \leq 1$ 情形 /120
 - 8.2.4 R^2 缺乏参考意义 /120
 - 8.3 LPM 以外的其他方法 /123
 - 8.4 Logit 模型 /124
 - 8.4.1 Logit 模型的估计: 个体数据 /124
 - 8.4.2 Logit 模型的估计: 分组数据 /126
 - 8.5 Probit 模型 /130
 - 8.6 三类模型比较 /131
 - 本章小结 /132
 - 练习题 /132
- 第 9 章 面板数据模型初步 /135
- 9.1 引入面板数据的背景 /135
 - 9.1.1 什么是面板数据 /135
 - 9.1.2 混合 OLS /136
 - 9.1.3 不可观测的异质性 /136
 - 9.2 固定效应模型 /137
 - 9.3 随机效应模型 /139
 - 9.4 Hausman 检验 /140
 - 本章小结 /141
 - 练习题 /141
- 附录 A /143
- 参考文献 /148

学习目标

1. 了解计量经济学的发展历史
2. 了解计量经济学的基本概念
3. 掌握计量经济学建模的一般步骤

计量经济学作为经济学专业的三大核心课程之一，几乎与宏观经济学同时出现在历史舞台上，其发展历时不足百年。20 世纪 30 年代，以挪威人弗里希为代表的一批经济学家，在面对电气化时代所出现的大批量经济观测数据时，深感有必要建立一套有力的分析工具，以便有效整理、分析在当时看来过于庞大的经济数据。从那时起计量经济学就以一个独立的分支开始出现在经济学的研究领域。经过近一个世纪的快速发展，计量经济学的工具箱已非百年前所能比，分析的方法更为精细，模型的种类更为丰富，所能处理的数据形式也更为复杂。尽管今天的计量经济学文献在面貌上与当初已有很大的区别，但是，其分析基础和研究目的则始终未变。不容置疑，计量经济学就是经济学的一个分支领域，它运用数理统计的某些方法对经济理论所假定的关系进行实证研究。

1.1 计量经济学发展历史

在学术界，一般认为“计量经济学”（econometrics）一词最早是由挪威人弗里希模仿生物计量学（biometrics）所提出的。在弗里希看来，如果经济理论在纯定性基础上“工作”，而不设法定量测度不同因素影响的重要性，是不可能得出和论证任何科学结论的。基于这种判断，通过以他为首的一批经济学家的大力倡导，世界计量经济学会（Econometric Society）于 1930 年 12 月 29 日成立，随后，在 1933 年，由该学会创办的学术刊物《计量经济学》（*Econometrica*）正式出版。这一系列事件标志着计量经济学作为一个独立的学科分支正式诞生了。正如萨缪尔森所说：“第二次世界大战（以下简称二战）后的经济学是计量经济学的时代。”在二战结束后，计量经济学迎来了第一次发展高峰。分析工具不断推陈出新，已有

的分析方法得到完善,新的分析工具不断被提出。受到宏观经济学中国民经济学运行体系的分析范式的启发,哈维尔莫(Haavelmo)、库普曼斯(Koopmans)以及克莱因(Klein)等人相继发表多篇关于联立方程的学术论文,把计量经济学的建模思路从单方程模型推进到多个方程(联立方程)模型,从而使研究者能更为系统地分析国民经济运行的内在规律。

在随后的半个多世纪中,由于众多计量经济学家的共同努力,计量经济学取得了进一步的发展。新问题、新数据、新方法不断涌现,使得计量经济学在时间序列分析、离散数据建模以及面板数据建模等方面取得重大突破,很多新的建模方法改变了人们传统的以回归分析为基础的建模框架。发端于20世纪60年代,成熟于90年代的非参数计量经济学则以一种全新的思路去研究变量之间的关系。进入新世纪以后,随着互联网技术以及数据存储技术的快速发展,大数据逐渐成为大家研究的热点,如何快速有效地分析大数据背后所隐含的量化,势必将触发计量经济学新一轮的大发展。

在中国,计量经济学的发展虽然起步晚,但速度并不慢。1960年,中国科学院成立了一个经济数学研究组,开始尝试用量化方法研究经济活动规律,但研究范围主要集中于投入产出、生产优化等领域。“文革”中,此研究一度停滞,改革开放以后,社科院于1979年成立了中国数量经济学会,并创办了一份学术杂志《数量经济技术经济研究》,正式与国际主流的研究范式接轨。1980年,中国数量经济学会首次举办计量经济学讲习班,因上课地点在颐和园龙王庙,又被称为龙王庙讲习班,这可以说是中国计量经济学的一次启蒙式教育,主讲教师由七位来华的美国专家组成,他们分别是:克莱因、邹至庄、刘遵义、萧政、粟庆雄、安德森、安藤。之后的十多年,计量经济学在中国的发展虽然遭受少许挫折,但整体发展方向未变。随着对外交流日趋频繁,越来越多的中国学者开始进入计量经济学的国际舞台,如艾春荣、白聚山、陈小红等人陆续在国际顶尖的学术期刊发表关于计量经济学理论方法方面的开创性论文,陈小红更是成为大陆首位世界计量经济学会会员(econometric society fellow)。

作为一门课程,计量经济学已经在我国高等院校的经管类专业有20多年的开设历史,其重要性也逐渐为人们所认识。1998年,教育部明确规定计量经济学为经济类专业的8门核心课程之一,目前,计量经济学课程已经在高校广泛开设,整体的教学水平、研究水平相比以往有了显著提高,学生学计量、用计量的意识也日渐加强。

1.2 什么是计量经济学

计量经济学在最初引进国内时,很长一段时间被译为经济计量学,因此从字面上解读就是经济测量的意思。虽然开展计量经济学的研究首先就是要对经济指标进行测量,但计量经济学有着更广泛的内涵。古扎拉蒂在其编著的《计量经济学基础》一书中罗列了七条不同学者给计量经济学所下的定义,我们摘录其中一部分如下:

萨缪尔森认为,计量经济学可定义为实际经济现象的数量分析。这种分析基于理论与观测的并行发展,而理论与观测又通过适当的推断方法得以联系。

戈登博格(Goldenberg)认为,计量经济学可定义为这样的社会科学:它把经济理

论、数学和统计推断作为工具，应用于经济现象的分析。

哈维尔莫认为，计量经济学的研究方法是，以统计推断的理论和技術作为桥头堡，以达到经济理论和实际测算相衔接的目的。

弗里希在《计量经济学》的创刊词中这样总结道：“经验表明，统计学、经济理论和数学这三者对于真正了解现代经济生活的数量关系都是必要的，但本身并非是充分条件。只有这三者结合起来才能形成力量，这种结合便构成了计量经济学。”

综合前人的观点，我们不难发现，从本质上来说，计量经济学就是一门经济学课程，它不是统计学，更不是数学。尽管在模型参数估计领域大量借鉴了统计学的方法和工具，但是如果离开了问题的经济背景、方法本身的经济学解释、方法应用的经济对象，计量经济学的众多方法不过是一堆“神秘”的数学符号而已。

1.3 计量经济学研究的一般步骤

传统的计量经济学进行实证研究时一般遵循如下步骤：①建立理论模型；②计量模型的设定；③数据的收集与整理；④模型的参数估计；⑤模型的检验；⑥模型的应用。建立模型包括根据相关理论假说建立理论模型，并以此为基础建立计量模型；数据的收集与整理工作是实证研究的基础，数据质量直接决定实证研究结果的质量，这一环节的工作通常在统计类课程中会被重点讨论；计量经济学教科书则主要论述后面三个步骤：模型的参数估计、模型的检验以及模型的应用等问题。

1.3.1 建立理论模型

计量经济学的实证研究总是关注经济领域的某个方面，对所要研究的经济问题，根据相关的经济理论确定经济指标之间的数学关系，即理论模型。经典的经济学理论或者数理经济学为计量经济学的实证研究提供了许多参考。例如，若想研究居民的消费行为特征，有经典的凯恩斯绝对收入假说，也有弗里德曼的持久收入假说，哪一个更适用？这正是计量实证要回答的问题；若想研究某个行业的企业生产规模是否适应企业所处的发展阶段，可以借鉴柯布一道格拉斯生产函数，通过分析劳动和资本的边际生产率所处的区间进行判断；若想分析中央银行的货币政策是否导致一国的通货膨胀，宏观经济学有经典的货币需求理论为我们的分析提供指导，费雪的货币方程式 $MV=PY$ ，是这一领域的经典理论，类似的例子不胜枚举。在 CC 学派^①的观点看来，一项有经济理论支撑的实证研究才是有现实意义的研究工作。

1.3.2 计量模型的设定

计量模型和经济理论模型或数理经济学的模型之间有联系，但也有区别。计量模型是

① CC 学派，或称 CC 方法论。它源自 1932 年美国商人 A. 考尔斯 (A. Cowles) 出资成立的一个基金委员会 (Cowles Committee)，用于资助计量经济学的研究。在该委员会的资助下，形成计量经济学研究中最为主流的范式，他们以概率统计为基础，强调经济理论对模型的指导意义。在不严格意义下，也称其为 CC 学派。

在经济理论模型基础上的继承与发展,理论模型描述的是经济变量之间的抽象联系,不涉及具体经济个体的行为差异,而且理论模型中的很多经济变量也是抽象的,在建立计量模型时,这些都需要具体化。在宏观经济学的教科书中,我们根据凯恩斯的理论给出了一个大家都非常熟悉的消费函数: $C=\alpha+\beta Y$, C 表示消费, Y 表示收入。这样一个纯数学模型对计量经济学家的帮助是有限的,当我们考虑对中国的城镇居民进行消费支出调查时,首先要确定消费、收入具体应该用什么指标来合理度量。被调查者一年中全部的开支能当作消费吗,是否应该剔除一些投资性支出?对于收入,也许用可支配收入更为合适。指标的合理设定都是计量建模者需要仔细思考的问题。假定我们抽取了100个家庭,那么这100个家庭的消费行为除了受到收入的影响外,还有可能受到其他一些因素的影响,比如家庭结构、户主的受教育程度、文化背景等。因此,在实际样本分析过程中,收入和消费绝不存在确定的一一对应关系,在构建计量模型时,我们考虑引入随机扰动项 ε_i 来表示那些不在理论模型中但同时又有可能影响被解释变量的影响因素。同时,为了区分不同的个体,我们给变量添加下脚标以示区别。修改之后的模型如下:

$$C_i = \alpha + \beta Y_i + \varepsilon_i \quad (1-1)$$

在这个计量模型中,等号左边的变量 C_i 被称为被解释变量,也称因变量; α 、 β 被称为参数, Y_i 被称为解释变量,或称自变量。

1.3.3 数据的收集与整理

样本数据的收集与整理,是建立计量经济学模型过程中最具挑战性的工作,也是对模型质量影响至关重要的一环。很多初学者可能会认为计量经济学的理论方法对他们而言是最大的障碍,但是实际上一旦掌握理论方法便不再有任何难度。然而实际问题千变万化,所涉及的数据也各不相同,很多时候理论模型很完美,但就是找不到合适的的数据,导致模型的估计无法顺利进行。对于有丰富实证经验的研究者而言,数据一定是他研究工作中可能遇到的最大障碍,正所谓“巧妇难为无米之炊!”

常规的获取数据的渠道主要有以下几种:政府机构的官方统计数据,如国家和省市地区的统计年鉴以及一些专业统计年鉴;研究机构的专业调查数据库,如中国健康与养老追踪调查 CHARLS 数据库、中国健康与营养调查 CHNS 数据库等;在金融领域,还有一些专业的数据库如 WIND、CSMAR、RESET、CCER 中国经济金融数据库等。当然,如果有经费支持,研究者也可自行针对所要研究的问题,开展问卷调查获取第一手数据资料。

收集所得的数据,按数据结构一般可以分为以下三类:时间序列数据、横截面数据和面板数据。时间序列数据是对同一个体某一指标持续一段时间的观测记录。如 GDP、CPI、失业率等指标的年度或季度数据,均为时间序列数据。横截面数据是对多个个体的某一指标在某一时刻的观测记录,如全班同学的身高或体重的数据。面板数据是对多个个体的某一指标持续一段时间的观测记录,如全国各个省市自治区过去5年的GDP统计数据。

数据的整理是保证数据质量很重要的一项工作。根据建模的需要,有时我们需要将一

些名义变量的指标进行调整,使之变为实际变量。有些数据中的个别指标可能会出现数据缺失,这时还需要采用特定的技术将缺失数据补充完整。

1.3.4 模型的参数估计

有了观测数据之后,接下来就是估计模型中的参数。从技术上讲,模型的参数估计包括模型的识别和估计方法的选择两个方面,实际操作上则主要依赖于软件。在经典的计量经济学中,参数估计主要有两种方法:最小二乘法 and 极大似然估计,当然还有广义矩估计方法,但它的技术性细节远超出本书所设定的难度,因此不在讨论之列。在本书接下来的几章当中,将重点讨论最小二乘法以及它在某些特定场合下的拓展,后面我们也将定性响应模型部分展示极大似然估计的应用。

1.3.5 模型的检验

在模型参数估计以后,我们初步得到一个“可见”的计量模型,这样的模型是不是现实中的一个较好的近似,还必须通过适当的规则来判断估计结果是否与理论预期一致。正如弗里德曼的观点,凡是不能通过经验证据来证实的理论或假设,都不能称为科学探索的一部分。对于一个计量模型而言,一般有三个层面的检验,即经济意义检验、统计学检验和计量经济学检验。

经济意义检验主要是检验模型的参数估计值是否有合理的经济意义,主要是考察参数估计值在符号、取值大小方面与预先的理论预期是否一致。有些模型中的参数,在理论上有着确定的取值区间,如果估计结果超出这一区间,则无法对结果进行合理的解释,这种情形必须重新构建模型。例如前面讨论消费函数,系数 β 表示的边际消费倾向,其取值应在0到1之间。如果估计结果超出这一范围,显然不合理,这样的估计结果是不能作为最终报告的。

统计学检验是对模型的参数估计结果,依照统计推断的标准对模型的显著性进行检验,包括系数的显著性检验和方程的显著性。前者主要是判断从统计意义上看,解释变量是否对被解释变量存在影响,后者则是判断模型的方程结构设定是否合理。

在计量模型估计之前,通常会对模型的结构以及所使用的数据做出一些假设,以保证估计结果的合理性。计量经济学检验主要是判断所用数据是否满足模型所要求的一些基本假定。

1.3.6 模型的应用

当所估计的模型通过了一些检验之后,我们就可以用这个模型来进一步研究和分析相应的经济问题。首先我们可以对经济变量之间的关系做出度量,从模型的系数中发现经济变量的结构关系,给出政策评价的一些量化依据。进行经济预测是模型的另一个重要应用,可以说,在计量经济学的早期,预测是计量模型最为重要的应用,只是最近这些年,计量模型应用才开始变得多元化,如政策评价、结构分析等。但对实际工作者而言,预测

还是他们所能想到最常用的建模功能。

1.4 软件的使用

对目前的计量经济学的学习，仅掌握一些理论工具是远远不够的。要想从事实际的建模工作，必须要掌握一款计量软件。由于计量经济学的分析方法很多都是源自统计学，因此很多统计软件也能处理计量模型。本书配套的软件是 EViews，当然类似如 SPSS、SAS、STATA 以及时下非常流行的 R 软件都能胜任计量经济学模型的参数估计工作。

相比其他数据处理软件而言，EViews 软件的功能相对集中，与计量经济学经典方法和模型匹配更好。在最新的版本中，EViews 软件能有效处理各种类型的数据，例如截面数据、时间序列数据和面板数据，都有与之对应的功能模块。可以说，EViews 软件是一款小而精的计量分析软件，对于它，初学者能快速上手。对于处理复杂数据的研究人员而言，它也能胜任绝大部分工作。当然，对于一些从事理论方法创新的研究者而言，也许 R 软件更适合。

本章小结

1. 计量经济学的出现已有近百年的历史，在中国也经历了 30 多年的发展，无论在理论方面还是在实践应用方面都取得了丰硕的成果。
2. 不同的学者对计量经济学的定义表述可能不尽相同，但都在强调计量经济学作为一门经济学课程的学科属性。
3. 计量经济学作为一门强调学以致用用的课程，在进行实证研究时有一套基本研究流程，包括建立模型、数据的收集与整理、模型的参数估计、模型的检验以及模型的应用。
4. 选择一款合适的计量软件是学好、用好计量经济学的关键。

练习题

1. 计量经济学在经济研究中的作用是什么？
2. 计量模型在对经济变量的数量关系进行探索的同时，能否说明变量之间是否存在因果关系？
3. 列举出几类常用的数据收集渠道。
4. 理论模型与计量模型的主要区别是什么？

统计基础知识回顾

学习目标

1. 了解概率、随机变量和概率密度等基本概念
2. 了解多元随机变量中联合密度、边缘密度、条件密度等概念
3. 掌握期望、方差、协方差和相关系数的定义及其运算规则
4. 熟悉几类常见的统计分布
5. 掌握参数估计的基本思想及其构造方法
6. 掌握假设检验的基本思想及其检验的一般步骤

2.1 随机变量与概率分布

2.1.1 概率、样本空间和随机变量

1. 概率和结果

当你登上一辆公交车的时候里面已有的乘客人数、期末考试的成绩、某个交易日上证综指的收盘点数、今天上午你接到的电话次数等这些事件的结果都含有偶然性或随机性因素。在这些例子中你都不知道最后的结果是什么。一个更为严谨的表述是，一个随机试验可能发生的互斥的后果被称为**结果**。如上午接到电话的次数可能只有一次，也可能有五次，又或者一次都没有，这些结果显然是互斥的，即只有一种后果会发生，当然，每一种结果发生的可能性不尽相同。

结果出现的概率反映的是在长期观测过程中，特定结果出现的比例。这是对概率一个不严谨但是比较直观的认识。如果去营业厅打印你过去一个月的通话记录，30天中有8天的上午一个电话都没接到，那么你可以认为，在某天上午你一个电话也没有接到的概率是 $4/15$ 。

2. 样本空间和事件

所有可能结果的集合被称为**样本空间**，而**事件**则指样本空间中的子集，即事件是一个

或多个结果的集合。事件——“今天上午接到的电话不超过 2 个”是包含没有接到电话、接到一个和接到两个电话的集合。

3. 随机变量

随机变量是随机试验发生结果的观测记录，这种观测记录可以是一些离散数值，也可以是连续区间上的数值。如果随机变量取值是离散数值，如上午接听电话的次数，只能是 0, 1, 2, …, 则这类随机变量被定义为离散型随机变量。而像上证综指的收盘点数可能会在某一区间上取任意值，因此可视为连续型随机变量。

2.1.2 离散型随机变量的概率分布

1. 概率分布

离散型随机变量的概率分布为变量的所有可能取值及其对应发生概率的列表，且所有的概率相加之和为 1。例如 X 表示上午接听电话的次数，合理的取值范围是所有非负整数。记 X 所有可能的取值为 k ($k=0, 1, 2, \dots$)，则 X 取各个可能值的概率表示如下：

$$P(X = k) = p_k \quad (2-1)$$

由概率的定义可知， P_k 应满足下面两个条件：

$$P_k \geq 0, k = 0, 1, 2, \dots \quad (2-2)$$

$$\sum_{k=0}^{+\infty} P_k = 1 \quad (2-3)$$

式 (2-1) 通常被称为离散型随机变量 X 的概率分布或分布率。

2. 贝努利分布

结果取 0 或 1 的二值随机变量是一类很常见的离散随机变量，称为贝努利随机变量 (Bernoulli random variable)，其对应的概率分布则称为贝努利分布。例如 X 表示在人群中随意找出一个人的性别，若 $X=0$ 表示女性， $X=1$ 表示男性，则 X 的概率分布可表示为：

$$X = \begin{cases} 1, & \text{概率为 } p \\ 0, & \text{概率为 } 1-p \end{cases} \quad (2-4)$$

式中， p 表示此人为男性的概率。式 (2-4) 表示的概率分布就是贝努利分布。

2.1.3 连续型随机变量的概率分布

1. 累积分布函数

连续型随机变量的累积分布函数为该随机变量小于或等于某个特定值的概率。对于任意实数 x ，随机变量 X 的分布函数 $F(x)$ 可定义为：

$$F(x) = P(X \leq x)$$

分布函数 $F(x)$ 具有以下性质：

1. $F(-\infty) = 0, F(\infty) = 1$
2. $F(x)$ 是一个单调非减函数，即如果 $x_1 < x_2$, $F(x_1) \leq F(x_2)$