

1

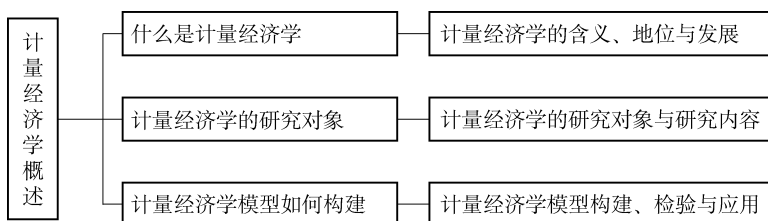
第一章

Chapter 1

计量经济学概述

>>>

知识结构图



>>>

学习目标

1. 知识目标：计量经济学的含义、地位与发展状况；计量经济学的研究对象与内容体系；计量经济学模型的含义、构建步骤、检验与应用。
2. 能力目标：理解计量经济学与数学、统计学、经济学之间的关系；理解计量经济学的内容体系；掌握计量经济学模型的构建步骤、检验方法与应用范围。

>>>

情景写实

妈妈说“多喝牛奶会长高”，但你发现邻居家孩子每天喝牛奶却没你高。

这是“遗漏变量偏差”问题。身高不仅受牛奶等饮食因素影响，还与遗传、运动、睡眠等因素相关。需要加入“控制变量”（父母身高、运动时长等）才能准确估计单一变量（牛奶）的真实影响。

你刷到的网红店评价全是好评，但自己去了却觉得一般。

这是“选择性样本偏差”问题。网红店的好评可能来自商家筛选的“托儿”，或只有满意的顾客才会主动评价。判断一件事时，别只看“显眼的数字”，要考虑样本是否全面，需要通过随机抽样避免样本偏差。

第一节 什么是计量经济学

计量经济学(Econometrics)是经济学的基础学科之一。它以一定的经济理论和统计资料为基础,运用数理方法与计算机技术等,以建立计量经济模型为主要手段,定量研究具有随机特性的经济变量,进而探究经济主体之间的互动规律。

理论经济计量学以计量经济学理论与方法技术为研究内容,目的在于为应用计量经济学提供方法论,主要研究如何运用、改造和发展数理统计方法,使之成为随机经济关系测定的特殊方法。

应用计量经济学是在一定的经济理论的指导下,以反映事实的统计数据为依据,用经济计量方法研究经济数量模型的实用化,或探索实证经济规律、分析经济现象和预测经济行为,以及对经济政策做定量评价。

相关链接 1-1

计量经济学的学科地位

计量经济学从诞生之日起,就显示出极强的生命力,而且经过20世纪四五十年代的大发展、60年代的大扩张、70年代现代计量经济学理论方法的研究与应用,已经在经济学科中占据非常重要的地位。正如著名计量经济学家、诺贝尔经济学奖获得者克莱因所评价的:“计量经济学已经在经济学科中居于最重要的地位。”“在大多数大学和学院中,计量经济学的讲授已经成为经济学课程表中最有权威的一部分。”著名经济学家、诺贝尔经济学奖获得者萨缪尔森甚至说:“第二次世界大战后的经济学是计量经济学的时代。”

计量经济学是经济学的一个分支学科,以揭示经济活动中客观存在的数量关系为主要内容。第一届诺贝尔经济学奖获得者、计量经济学的创始人、挪威经济学家拉格纳·弗里希(Ragnar Frisch)将计量经济学定义为经济学、统计学和数学三者的有机结合(图1-1)。

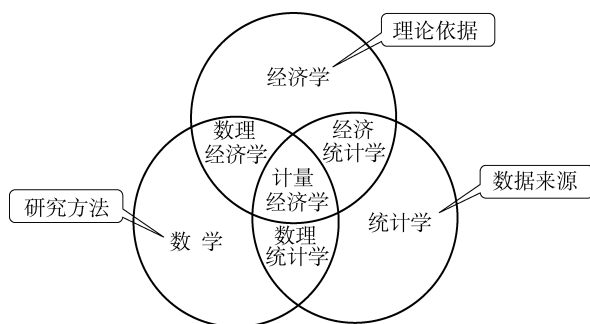


图 1-1 计量经济学的学科地位

20世纪70年代末,计量经济学在我国得到了迅速的传播与广泛的应用。目前,它对于经济管理工作的重要性不言而喻。

经典实例 1-1

计量经济学与经济学、数理经济学、经济统计学的学科内涵有什么异同?

经济学:研究如何最优配置和高效利用现有可供选择的有限资源,实现人类现在和将来无限欲望的最大满足。

数理经济学:研究如何运用抽象的方法得出经济学概念与理论,包括借助数学函数和几何图形等工具,不考虑对经济理论的度量和经验解释。

经济统计学:研究如何获得经济学统计资料和借助统计资料实现现实经济现象动态变化过程的高效记录,包括收集、整理、输出经济数据等。

计量经济学:研究如何以经济学统计资料作为数据来源,以经济学作为理论依据,以数学作为研究方法,以计量经济学模型作为主要手段,实现经济理论验证、未来经济预测和经济政策评价。

一、计量经济学的起源

数学方法在经济学中的应用,最早可追溯到 300 多年前,即 1676 年英国古典政治经济学的创始人威廉·配第(William Petty)编著的《政治算术》。“计量经济学”一词,是挪威经济学家弗里希在 1926 年仿照“生物计量学”一词提出的。随后不久,1930 年成立了世界计量经济学会,1933 年创办了在经济学界影响力非常大的《计量经济学》学术杂志。

对于人们应如何理解“计量经济学”,弗里希在《计量经济学》学术杂志的创刊词中解释:“用数学方法探讨经济学,可以从好几个方面着手,但任何一方面都不能与计量经济学混为一谈。计量经济学与经济统计学绝非一码事,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,经济学、统计学和数学这三者对于真正了解现代经济生活中的数量关系来说,都是必要的,但各自并非充分条件,而三者有机结合起来才有力量,这种有机结合便构成了计量经济学。”

相关链接 1-2

计量经济学的学科属性

任何一项科学研究,大体都需要经过以下几个阶段:首先是观察自然或社会现象,观察的现象通常都是偶然或随机的;然后设计严格的假设条件,在随机现象的基础上提出理论模型;最后对理论模型实行一系列严格的数理检验,如果数理检验通过了,说明理论模型是成立的。检验方法多样,工科、理科主要借助实验或实践,社会学可以借助各种数理统计检验。计量经济学中的回归模型就是理论模型,计量经济学正是通过各种数理检验(如 R^2 检验、 F 检验、 t 检验)验证回归模型成立与否的一门学科。

1998 年 7 月,教育部高等学校经济学学科教学指导委员会讨论并确定了高等学校经济学一级学科各专业的 8 门共同的核心课程,其中包括“计量经济学”。由于计量经济学教材中充满各种数学原理的应用,而当时对经济学等学科的数学基础认识不足,于是教育部将“计量经济学”纳入文科课程范畴,这也是当前许多高校出现计量经济学课程“教师难教、学

生难学”窘况的原因之一。

总体而言,要真正学好“计量经济学”,需要良好的数学、统计学和经济学知识基础,同时还需较强的思维分析能力,以及较专业的计量建模能力和计量软件(如 EViews 软件)的应用能力。对于基础不够好的学生,需要预习教材,多听老师讲授,多阅读国内外同类相关书籍,多进行计量软件的实训操作等。学习计量经济学,可为写作毕业论文和实证分析经济学现象等提供保障。

二、计量经济学的应用

目前,国内对计量经济学的理论研究很少,大多是关于计量经济学应用的研究。对现代计量经济学理论,国内基本上处于学习和跟踪阶段。计量经济学在经济管理等工作中的应用非常广泛,大致表现在以下几个方面。

(一) 理论检验

理论检验是计量经济学最主要的用途。从一系列样本观察开始,然后通过数理分析,概括出相关经济理论模型,这是理论构建。用已有的经济理论模型去拟合(或预测、应用等)现实世界,如果拟合得非常好,那么这一理论就得到检验;反之就未通过检验。这是理论检验。从逻辑上看,理论构建是一种归纳推理的过程,即从特殊到一般;而理论检验则是一种演绎推理的过程,即从一般到特殊。

社会经济研究的主要目的,是对社会经济现象及其规律(或理论模型)作出科学归纳和正确解释。一个成功的计量经济学模型,必须很好地拟合社会经济的样本数据。样本数据是已经发生了的经济活动,如果模型拟合得非常好,则模型中的数量关系就是经济活动所遵循的经济规律。

相关链接 1-3

冒险犯罪的经济模型

诺贝尔经济学奖得主加里·斯坦利·贝克尔(Gary Stanley Becker)设计了一个著名的计量经济学模型,试图验证个体犯罪行为的选择机制。通常每一项特定的犯罪都有较可观的经济回报,但当实施该犯罪行为时往往也有不可忽视的机会成本。根据利润最大化原理,理性的个体会在经济回报与机会成本的比较中作出选择,只有当经济回报大于机会成本时,个体才会冒险去犯罪。

计量经济学模型为 $y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7, \dots)$, y 为犯罪所获得的利润。其中: x_1 为犯罪活动获得的经济回报; x_2 为犯罪活动所耗时间; x_3 为个体单位时间的合法收入; x_4 为犯罪活动被抓的概率; x_5 为被抓后被判有罪的概率; x_6 为获罪后被监禁的时间长度; x_7 为个体年龄……

(二) 结构分析

当一个变量或几个变量发生变化时,会对其他变量以至整个经济系统产生一定的影响。结构分析就是对经济现象中变量之间关系的研究,如价格 P 与需求量 Q 之间的关系。结构分析所采用的主要方法有弹性系数法、边际分析法等。

相关链接 1-4

施肥量对大豆收成的影响

大豆的生长通常需要大量肥料,一般施肥量越大,大豆收成就越高。当然,施肥量只是影响收成的重要因素之一,还有土地肥力因素、降雨量、种群特征、耕作劳动量等。为了得到施肥量与收成之间的相互关系,可以选择几块面积相同、土地肥力相同的耕地,同时种上大豆。这几块耕地除了施肥量不同之外,其他影响要素都相同。分别对这几块耕地施撒不同倍数($n=1,2,3,4,5,\dots$)的肥料量,等收割后,登记这几块耕地的大豆收成,并根据结构分析法找出施肥量与大豆收成量之间的数量关系。

(三) 最优决策

最优决策是指决策者选择最优方案,追求理想条件下的最优目标。

决策行为包括以下几个要素。

- (1) 决策主体,通常分为个人决策者与集体决策者。
- (2) 决策目标,不同的决策其目标函数可能不同。
- (3) 决策变量,指影响决策目标并且决策者能够调控的因素。
- (4) 决策参数,指影响决策目标而决策者无法调控的因素。
- (5) 约束条件,指决策变量与决策参数变化的范围及其相互关系,如社会经济因素、制度环境因素等。

按照决策的目标函数,其可以分为投入产出效率最大化、利润最大化、产量最大化、成本最小化等;按照决策变量,其可以分为行业组合决策、产品组合决策、产量决策、要素组合决策、投资组合决策、人力资源组合决策、营销组合决策等。

相关链接 1-5

粥分配的民主决策

从前,有9个伙伴住在一起,每天共喝一桶粥,粥每天都不够分。

一开始,他们采用抓阄的方式决定由谁来分粥。每次抽选一个人,几周下来,结果他们觉得只有自己分粥的那一次是饱的。

接着,他们想到采用投票的方式选出一个道德高尚的人当首长,让首长全权负责分粥。权力过分集中就易产生腐败,于是就有人挖空心思去讨好首长,甚至贿赂他。首长后来也越来越嚣张跋扈,顺我者粥多,逆我者粥少。直到某一天,该首长被举报家有大量粥来源不明并因而被免职。

然后,他们设计出三人的分粥立法委员会、三人的分粥行政委员会和三人的分粥司法委员会,各委员会相互监督。结果是三个委员会经常互相攻击,不断相互扯皮,办事效率低下,导致大家喝到嘴里的粥天天都是凉的。

最后,大家摸索出一个最优决策:实行轮流分粥制,但分粥的那个人必须要等其他人挑选完后,才能拿剩下的最后一碗。为了不让自己的粥最少,每个人每次分粥时都尽可能把粥分配平均。

自此以后,大家和和气气,日子越过越好。

(四) 政策评价

政策评价是研究不同的政策实施对经济目标所产生影响的差异。在采取某项政策前,有时先在局部范围内进行试验,然后推广与实行。

计量经济学模型,揭示了经济系统中变量之间的相互联系,将经济目标作为被解释变量,经济政策作为解释变量,可以很方便地评价各种不同的政策对目标的影响。一是预期目标法,给定目标变量的预期值,即希望达到的目标,通过求解模型,可以得到政策变量值;二是政策模拟法,将各种不同的政策代入模型,计算各自的目标值,然后比较其优劣,决定其政策的取舍。

相关链接 1-6

牛肉面馆老板的烦恼

在闹市入口有家牛肉面馆,一开始生意非常红火,但后来却不知什么原因不做了。

朋友问老板为什么,老板说:“现在的人贼着呢!我当时雇了个会做拉面的师傅,但在工资上总也谈不拢。开始的时候,为了调动师傅的积极性,我们按卖的多少来分成,每卖一碗面让他挣5毛钱。经过一段时间,他发现来吃面的客人越多,他的收入越高。这样一来,他就在每碗面里加超量的牛肉,吸引回头客。一碗面才4块钱,本来就靠薄利多销,他每碗多放几片牛肉,我还怎么挣钱?”

“后来我看这样不行,钱全被他赚去了!就换了个办法,给他每月发固定工资,工资给高点也无所谓,这样他不至于多加牛肉了吧?因为客多客少和他的收入没关系。但你猜怎么着?”老板有点儿激动了,“他在每碗里都少放牛肉,把客人都赶走了!”“这又是为什么?”朋友激动地问。“牛肉的分量少,顾客就不满意,回头客就少,生意肯定清淡,他才不管你赚钱不赚钱呢,反正他拿的是固定的工钱,卖多少无所谓,没客人他才清闲呢!”

(五) 经济预测

经济预测是与未来有关的旨在减小不确定性对经济活动影响的一种经济分析,不是靠经验、凭直觉或猜测,而是以科学的理论方法、可靠的资料、精密的计算及对客观规律性的认识作出的分析和判断。

计量经济学模型作为一类经济数学模型,是以模拟历史、从已经发生的经济活动中找出变化规律的技术手段,从经济预测特别是短期预测发展起来的。在20世纪五六十年代,运用计量经济学模型对西方国家经济预测不乏成功的实例。

经典实例 1-2

计量经济学模型能告诉我们什么

以某区域汽车市场计量经济学模型为例, $y=f(x_1, x_2, x_3, x_4, x_5, x_6, x_7, \dots)$, y 为汽车销售量,其中: x_1 为该区域人均收入水平; x_2 为汽车价格; x_3 为汽油价格; x_4 为汽车售后服务水平; x_5 为交通罚款平均值; x_6 为汽车寿命; x_7 为银行利率水平……

一旦该区域的汽车市场计量经济学模型各参数计算出来,即完整的计量经济学模型已

经建立。那么,我们可以获得以下信息:

- (1) 影响汽车销售量的主要因素是什么(收入、价格等)?
- (2) 各种因素对汽车销售量影响的性质怎样(正、负)?
- (3) 各种因素影响汽车销售量的具体数量程度怎样(各变量的参数值)?
- (4) 以上分析所得结论是否可靠(F 检验、 R^2 检验、 t 检验等)?
- (5) 今后发展的趋势怎样(经济预测)?

三、计量经济学的发展

1969 年,首届诺贝尔经济学奖授予了“计量经济学奠基人”挪威经济学家弗里希和“计量经济模式建造者之父”荷兰经济学家简·丁伯根(Jan Tinbergen)。随后,一半以上的诺贝尔经济学奖颁给了对计量经济学模型颇有建树的经济学家。

近些年来,诺贝尔经济学奖两次授予计量经济学的分支学科,2000 年是表彰詹姆斯·J. 赫克曼(James J. Heckman)和丹尼尔·L. 麦克法登(Daniel L. McFadden)对横截面数据的分析方法作出的杰出贡献,2003 年是表彰罗伯特·F. 恩格尔(Robert F. Engle)和克莱夫·W. J. 格兰杰(Clive W. J. Granger)分别用“随着时间变化的易变性”和“共同趋势”两种新方法分析经济时间数列。这都说明,计量经济学“技术层面”的研究工作越来越得到广泛的认可和高度的重视。

计量经济学早已不再是数理统计在经济学领域的简单应用,针对研究对象的独特性,计量经济学不断开创出新的理论和方法,如广义矩估计、协整分析、因果关系、脉冲效应、高频数据处理、虚拟变量处理、面板数据处理、状态空间模型、联立方程模型等,不仅拓展和丰富了统计学的研究范式,也给人们带来了对外经济社会的全新认识。

如今,计量经济学与微观经济学、宏观经济学一起构成了现代经济学的三大核心。计量经济学对社会学、政治学乃至历史学等的研究也产生了深远的影响。它们也越来越多地借鉴和使用计量经济学的分析工具。在我国,计量经济学经过 30 多年的发展,计量经济学模型已经成为经济理论研究和实际经济分析的主流实证方法。当前,科研论文如果没有计量经济学模型实证分析,很难发表在《美国经济评论》和《经济研究》《中国社会科学》《管理世界》等国内外顶级学术期刊上。

一项实证研究从计量经济学模型的设定开始,一直到模型的估计、检验、评价和解释,其随意和错误随处可见。于是,人们对计量经济学模型方法产生了不同的甚至是相反的评价,究其原因:部分来自计量经济学模型方法本身,更多来自计量经济学模型的应用研究。国内学者李子奈、潘文卿、高铁梅、叶阿忠、张晓峒、于俊年等为我国计量经济学的普及作出了奠基性的贡献。

相关链接 1-7

三次著名的计量经济学国际大讨论

计量经济学作为一门独立的经济学分支学科,其区别于其他相关学科的本质特征是什么?计量经济学应用研究的科学性和可靠性如何保证?这些问题引发了三次著名的计量经济学国际大讨论。

第一次大讨论,始于有名的“凯恩斯-丁伯根之争”(Keynes, 1939、1940; Tinbergen, 1940),凯恩斯认为丁伯根所用的多元回归分析是一种“巫术”,计量经济学作为“统计炼金术”的分支还远未成熟到足以成为科学的分支。凯恩斯反对使用概率论,而丁伯根的“回归”却未能利用概率论的原理很好地解释估计结果,当时的经济学实证研究陷入困难丛生的境地。最后这场争论以 Haavelmo(1944)《计量经济学中的概率论方法》一文的发表而告结束。该文为经济学中的概率论原理正名,并在概率论的基础上建立起统一的计量经济学基本框架。

20 世纪 80 年代初,掀起了有关经验研究可信性问题的第二次大讨论。Sims(1980)对当时大型宏观计量经济模型的外部约束条件的可靠性提出质疑,认为这些不现实的约束条件将导致不可靠的政策分析结论,建议使用更少约束条件的 VAR 建模策略。该模型已被研究者和政策制定者所广泛采用,主要用于分析经济如何受到经济政策临时性变化和其他因素的影响,Sims 也因此获得 2011 年诺贝尔经济学奖。Hendry(1980)对计量经济学沦为“炼金术”问题展开了尖锐的批判,提出经验研究走向科学的一条金科玉律——“检验、检验、再检验”。Leamer(1983)指出模型假定以及控制变量选择的随意性会导致结果的脆弱性,提倡进行回归模型的敏感性分析。Black(1982)以及 Pratt 和 Schlaifer(1984)对随机变量之间的相关关系错误推广至因果关系等现象提出了批判,同时对两者的区别进行了详细的论述。随后,计量经济学家提出了各种建模思想、估计量以及检验统计量,理论计量进入百花齐放的阶段。

然而,理论计量研究与经验研究之间的裂缝反而扩大了,理论计量越来越复杂,应用计量则在某些领域变得越来越简单(Heckman, 2001)。为此,进入 21 世纪以来,以 *Journal of Econometrics* 百期纪念专刊对计量经济学方法论、模型方法发展的总结为开端,以重要学术期刊的专刊为阵地,计量经济学界掀起了对经验研究可信性的第三次大讨论。一场经验研究的“可信性革命”(Angrist and Pischke, 2010)蔚然成风,并形成了模型设定的统计适切性和因果关系的有效识别两大核心议题。

综观三次大讨论,可信性革命的核心问题在于实现经济学、统计学、数学在计量经济学应用研究中的科学结合。第一次大讨论主要关注经济理论与数学的结合问题,解决了计量经济学的概率论基础问题,同时确立了凯恩斯宏观经济理论在模型设定中的导向作用。第二次大讨论突出了数据与模型的结合问题,在宏观实证领域摒弃了模型设定的经济理论导向,确立了数据关系的导向作用。第三次大讨论强调了模型设定的统计适切性问题和因果关系的有效识别问题,本质上是试图实现经济理论导向和数据关系导向的综合,向实现经济理论、统计学、数学的科学结合迈出了坚实的一步。

第二节 计量经济学的研究对象

计量经济学主要有两大研究对象:横截面数据(Cross Sectional Data)和时间序列数据(Time Series Data)。前者旨在归纳不同经济行为者是否具有相似的行为关联性,以模型参数估计结果显现相关性;后者重点在分析同一经济行为者不同时间的资料,以展现研究对

象的动态行为。

新兴计量经济学研究切入同时具有横截面及时间序列的资料,换言之,每个横截面都同时具有时间序列的观测值。这种资料称为面板数据(Panel Data)。面板数据研究多个不同经济体动态行为之差异,可以获得较单纯横截面或时间序列分析更丰富的实证结论。例如诺贝尔经济学奖得主格兰杰指出,在回归模型中对一组检验进行诠释进而揭示因果关系是可行的,其所提出的格兰杰因果关系检验强化了回归分析(Regression Analysis)的说服力。

经典实例 1-3

当代计量经济学的主要研究领域

当代计量经济学研究,主要集中在单位根检验、时间序列模型、波动模型、向量自回归模型与向量误差修正模型、离散选择模型与受限模型和面板数据模型六个领域。其他研究领域还有非参数与半参数估计、广义矩估计、贝叶斯估计、分数积分研究、模拟与自举技术等。

一、横截面数据

横截面数据是指在同一时间(时期或时点)截面上反映一个总体的一批(或全部)个体的同一特征变量的观测值,是样本数据中的常见类型之一。横截面数据是由同一时间、不同统计单位、相同统计指标组成的数列。与时间序列数据相比较,其区别在于组成数据列的各个数据的排列标准不同,时间序列数据是按时间顺序排列的,横截面数据是按照统计单位排列的。例如,工业普查数据、人口普查数据、家庭收入调查数据。

横截面数据不要求统计对象及其范围相同,但要求统计的时间相同,即必须是同一时间截面上的数据。与时间序列数据完全一样,横截面数据的统计口径和计算方法(如价值量的计算标准)也应当是可比的。例如,为了研究某一行业各个企业的产出与投入的关系,需要关于同一时间截面上各个企业的产出 Q 和劳动投入 L 、资本投入 K 的横截面数据。这些数据的统计对象显然是不同的,因为是不同企业的数据,但是关于产出 Q 和投入 L 、 K 的解释、统计口径和计算方法仍然要求相同,即各企业的 Q 、 L 、 K 在统计上要求可比。

分析横截面数据时,应主要注意两个问题:一是异方差问题,由于数据是在某一时期对个体或地域的样本的采集,不同个体或地域本身就存在差异;二是数据的一致性,主要包括变量的样本容量是否一致、样本的取样时期是否一致、数据的统计标准是否一致。

二、时间序列数据

时间序列数据是在不同时间点上收集到的数据。这类数据反映某一事物随时间的动态变化状态,如2004—2024年的国内生产总值(GDP)。时间序列数据除了年度数据之外,还有季度数据、月度数据、周度数据、天度数据、时度数据等。

时间序列分析是根据系统观测得到的时间序列数据通过曲线拟合和参数估计等建立数学模型的数量方法。其基本原理:一是承认事物发展的延续性,应用过去数据,推测事物的发展趋势;二是考虑到事物发展的随机性,因为任何事物的发展都可能受偶然因素影响,为此要利用加权平均法对历史数据进行处理。该方法常应用在国民经济宏观控制、区域综合发展规划、企业经营管理、市场潜量预测、气象预报、环境污染控制等方面。

一个时间序列通常由趋势、季节变动、循环波动和不规则波动四种要素组成。趋势是时间序列在长时期内呈现出来的持续向上或持续向下的变动；季节变动是时间序列在一年内重复出现的周期性波动，是诸如气候条件、生产条件、节假日或风俗习惯等因素影响的结果；循环波动是时间序列呈现出非固定长度的周期性变动，是涨落相同的交替振动；不规则波动是时间序列中除去趋势、季节变动和循环波动之后的随机波动。

不规则波动通常总是夹杂在时间序列中，致使时间序列产生一种波浪形或震荡式的变动。只含有不规则波动的序列也称为平稳序列。

相关链接 1-8

时间序列数据的输入与编辑

EViews 8.0 软件操作步骤：

(1) 创建工作文件。打开 EViews 工作界面，单击 File/New/Workfile，如图 1-2 所示。



图 1-2 新建工作文件的命令

(2) 完成 Workfile Create。选择 Workfile structure type(文件结构类型)为 Dated-regular frequency(时间序列)，选择 Date specification(时间标准)中 Frequency(频率)为 Annual(年度)，Start date(开始年份)如 2000 年，End date(结束年份)如 2025 年，单击 OK 按钮，如图 1-3 所示。

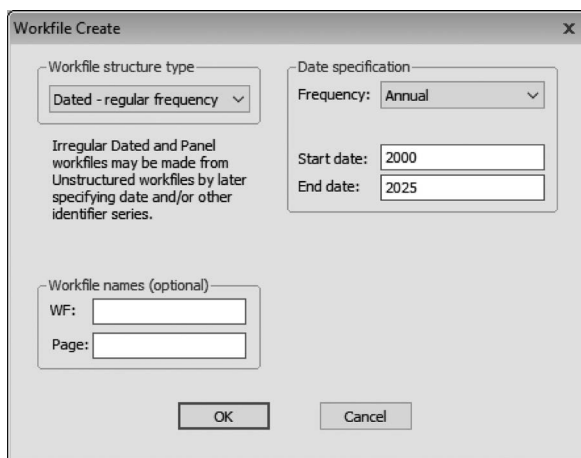


图 1-3 新建工作文件的对话框

出现 c(常数项)和 resid(残差项)两个数据系列(图 1-4)。

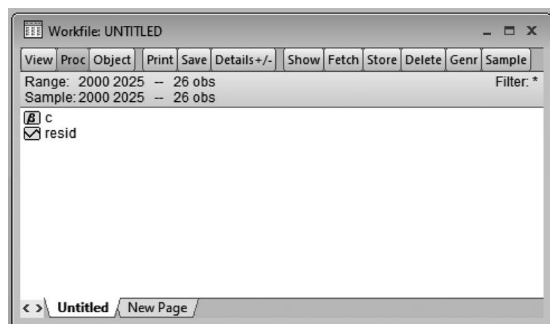


图 1-4 常数项和残差项系列

(3) 创建工作组。单击最上层标题栏 Quick, 进入 Empty Group(空组), 如图 1-5 所示。

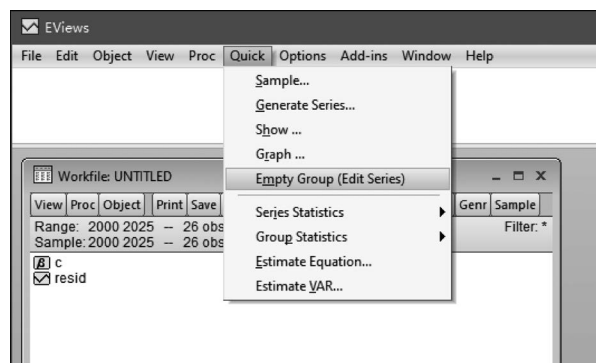


图 1-5 新建工作组的命令

出现 Group 组变量与数据输入界面(图 1-6)。

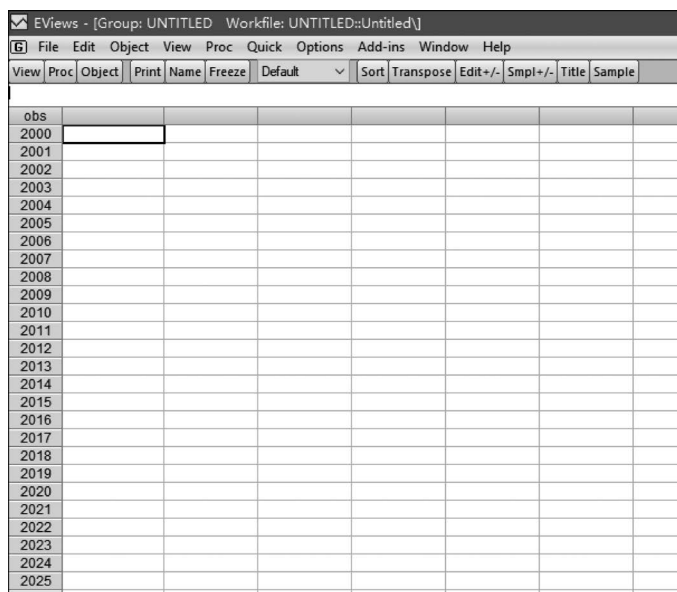


图 1-6 数据输入界面

(4) 输入变量名与变量对应的样本数据。在 obs(观察值)中输入 Y 与 x1,x2 等变量名,将光标移出格子外,单击,得到 Series create 界面,如图 1-7 所示;可以选择变量的属性,如 Numeric series(数值系列)。

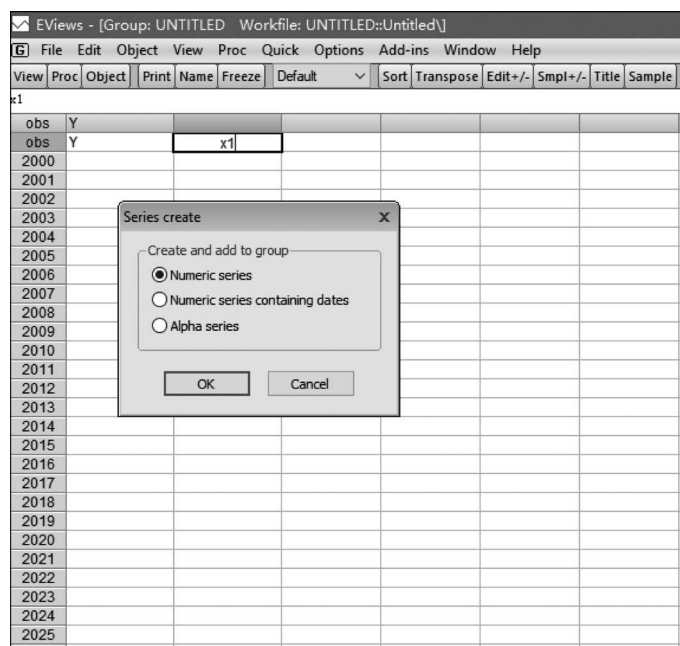


图 1-7 变量属性界面

在 NA(数值)时间系列中输入样本数据,如图 1-8 所示。

EViews - [Group: UNTITLED Workfile: UNTITLED::Untitled]					
File Edit Object View Proc Quick Options Add-ins Window Help					
View	Proc	Object	Print	Name	Freeze
Default	Sort	Transpose	Edit+/-	Smpl+/-	Title
Sample					
obs	Y				
obs	Y	x1			
2000					
2001					
2002					
2003					
2004					
2005					
2006					
2007					
2008					
2009					
2010					
2011					
2012					
2013					
2014					
2015					
2016					
2017					
2018					
2019					
2020					
2021					
2022					
2023					
2024					
2025					

图 1-8 变量样本数据输入

在文件界面自动增加了 y, x1, x2 数据系列, 如图 1-9 所示。

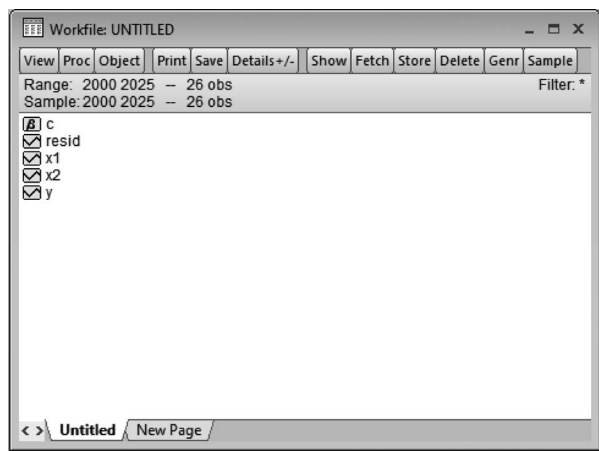


图 1-9 变量数据系列

三、面板数据

面板数据有时间序列和横截面两个维度。当这类数据按两个维度排列时, 是排在一个平面上, 与只有一个维度的数据排在一条线上有着明显的不同, 整个表格像是一个面板。例如: A、B、C、D 四个乡镇 2025 年的地区生产总值分别为 8、9、10、11(亿元), 这是横截面数据; A 乡镇 2025、2026、2027、2028 四年的地区生产总值分别为 8、9、10、11(亿元), 这就是时间序列数据; A、B、C、D 四个乡镇 2025、2026、2027、2028 四年的地区生产总值(亿元)分别为

A 乡镇(2025—2028): 8、9、10、11

B 乡镇(2025—2028): 9、10、11、12

C 乡镇(2025—2028): 10、11、12、13

D 乡镇(2025—2028): 11、12、13、14

这就是面板数据。面板数据可以克服时间序列数据多重共线的困扰, 能够提供更多的信息、更多的变化、更多的自由度和更高的估计效率。

相关链接 1-9

面板数据的输入与编辑

EViews 8.0 软件操作步骤:

(1) 创建工作文件。打开 EViews 工作界面, 单击 File/New/Workfile。

(2) 完成 Workfile Create。选择 Workfile structure type(文件结构类型)为 Balanced Panel(平衡面板), 选择 Panel specification(面板标准)中 Frequency(频率)为 Annual(年度), Start date(开始年份)如 2000 年, End date(结束年份)如 2025 年, Number of cross sections 中输入截面数据成员个数, 如图 1-10 所示, 单击 OK 按钮。

出现 c(常数项)和 resid(残差项)、crossid(截面项)、dateid(时间序列项)四个数据系列, 如图 1-11 所示。

(3) 创建 Pool(数据库)对象。单击最上层标题栏 Object, 进入 New Object(新对象), 如图 1-12 所示。

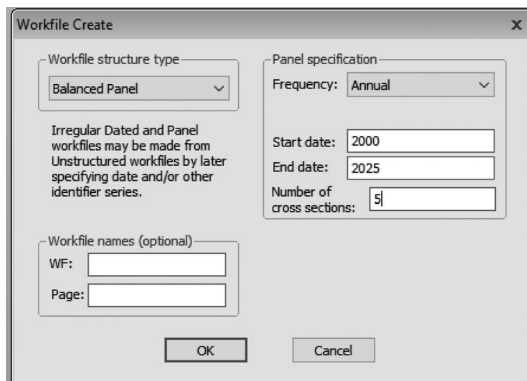


图 1-10 面板数据属性设置

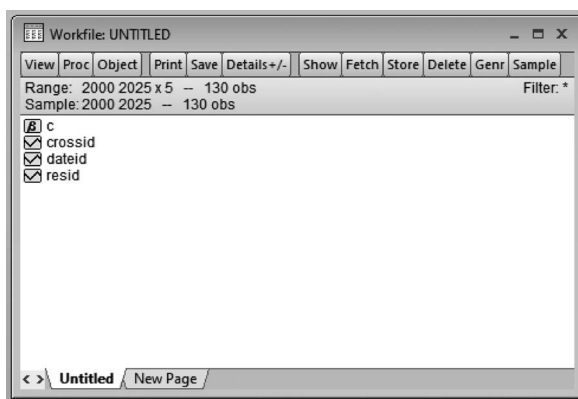


图 1-11 变量数据系列

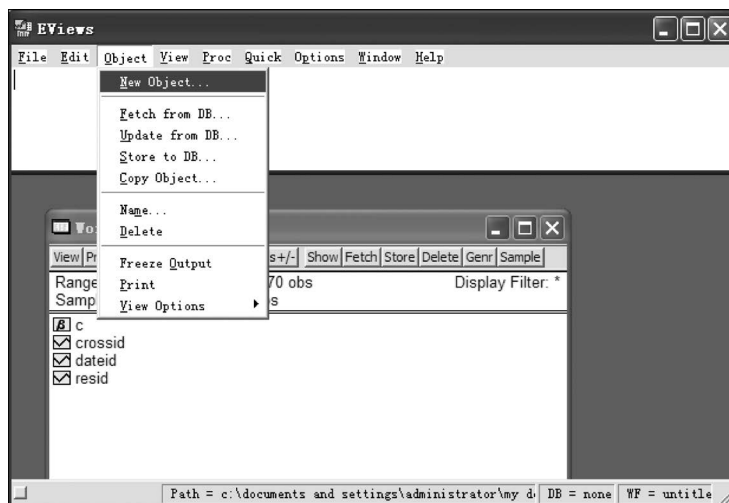


图 1-12 建立新对象的命令

出现 New Object 界面,在 Type of object(对象类型)中选择 Pool(数据库),在 Name for object(对象名)中输入对象名称,单击 OK 按钮,如图 1-13 所示。

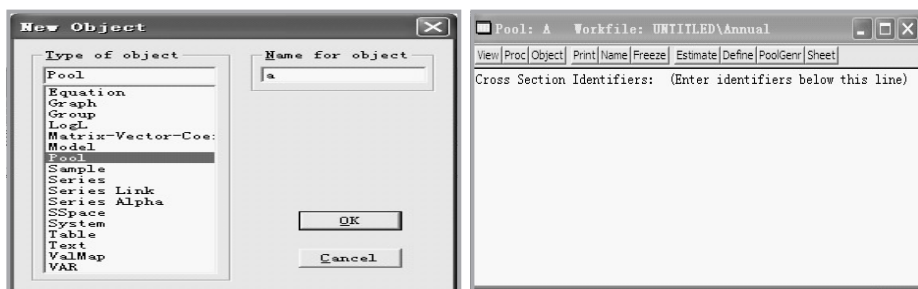


图 1-13 建立 Pool 对象

出现 Pool 的编辑窗口,可以在窗口输入截面成员的标识名。例如中国省际面板数据,可以选取湖南省、湖北省、河南省、江西省、安徽省中部五省份,分别用字母 HN、HB、HN、JX、AH 表示,一般每一个标识名占一行,如图 1-14 所示。

(4) 输入 Pool 对象的样本数据。打开 Pool 对象,选择 View/Spreadsheet (stacked data) 电子表格菜单命令,如图 1-15 所示;在弹出的对话框里输入序列名称,如 Pool01,如图 1-16 所示,单击 OK 按钮;出现 Pool01 数据表,单击工具栏中的 Edit+/- 可对样本数据进行输入和编辑,如图 1-17 所示。

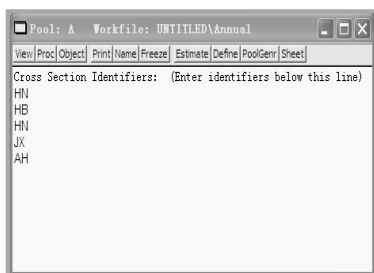


图 1-14 Pool 编辑窗口

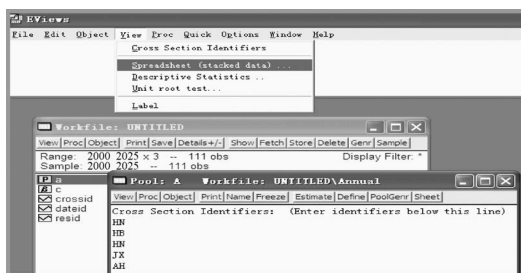


图 1-15 Pool 电子表格



图 1-16 Pool 对象序列名称输入



图 1-17 Pool 对象样本数据输入

第三节 计量经济学模型如何构建

一、什么是计量经济学模型

模型是用人们所熟悉的形式(几何或数学或实物)去解读和说明现实世界的某些功能(或

规律)。模型是所研究的系统(过程或事物)的一种表达形式,是对现实世界的抽象,如图 1-18 所示。模型是人类思维的创造物,通常易于操作和演练,具有重要的仿真模拟功能。

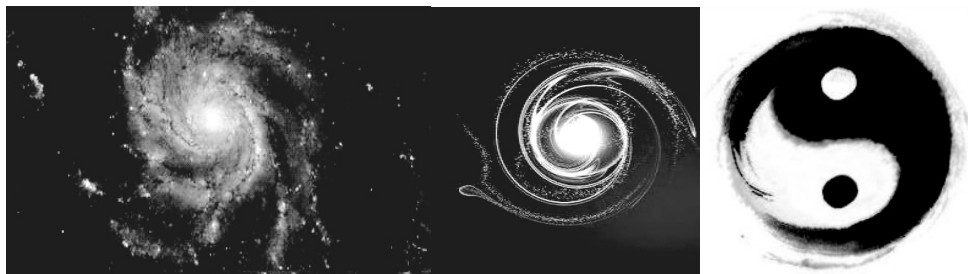


图 1-18 星系与太极图

计量经济模型(The model of Econometrics)是表示经济现象及其主要因素之间数量关系的方程式,通常用随机性的数学方程加以描述。数学方程式主要由经济变量、参数以及随机误差三大要素组成。例如道格拉斯生产函数:

$$Q = Ae^{\gamma t} K^{\alpha} L^{\beta} \mu \quad (1-1)$$

式中, Q 、 K 、 L 、 t 为经济变量; A 、 γ 、 α 、 β 为参数; μ 为随机误差项。

相关链接 1-10

大豆收成与施肥量的计量经济学模型

假定大豆收成的计量经济学模型为: $Y = \alpha + \beta X + \mu$,其中 Y 为大豆收成; X 为施肥量; α 、 β 为参数; μ 为随机误差项。农业研究者感兴趣的是:在其他因素不变的情况下,施肥量如何影响大豆收成。参数 β 给出了这个答案,因为 $\Delta Y = \beta \Delta X$ 。当然,施肥量只是影响收成的重要因素之一,随机误差项 μ 包括土地肥力因素、降雨量、种群特征、耕作劳动量等因素。

经济变量(Variable)是反映经济变动情况的随机变量,分为自变量和因变量,如式(1-1)中的 Q 为因变量, K 与 L 、 t 为自变量。为了描述自变量与因变量的变化,引入一些不是当前问题必须研究的另外变量,这样的变量叫作参数(Coefficient),一般反映事物之间相对稳定的比例关系。

自变量也可分为内生变量和外生变量。内生变量是在经济体系内部由纯粹经济因素影响而自行变化的变量(通常不被政策因素所左右),如式(1-1)中的 K 、 L 、 t ;外生变量是在经济机制中受外部因素(主要是政策因素)影响而非经济体系内部因素所决定的变量,如式(1-1)中的技术指数 A 。

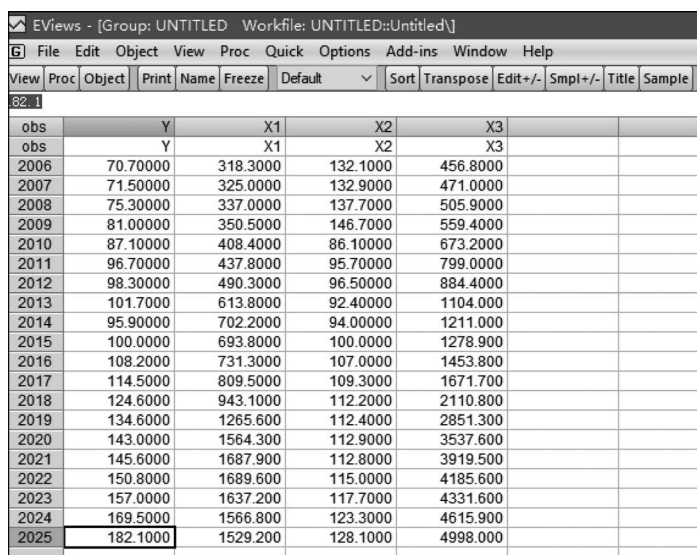
随机误差项(Stochastic Error)是指那些很难预知的偶然性误差,包括统计与整理经济资料、模型运算等过程中所出现的差错,通常正负误差可以抵消,可忽略不计。

相关链接 1-11

经济变量的统计描述

EViews 8.0 软件操作步骤:

(1) 输入变量名与样本数据,如图 1-19 所示。



obs	Y	X1	X2	X3
2006	70.70000	318.3000	132.1000	456.8000
2007	71.50000	325.0000	132.9000	471.0000
2008	75.30000	337.0000	137.7000	505.9000
2009	81.00000	350.5000	146.7000	559.4000
2010	87.10000	408.4000	86.10000	673.2000
2011	96.70000	437.8000	95.70000	799.0000
2012	98.30000	490.3000	96.50000	884.4000
2013	101.7000	613.8000	92.40000	1104.000
2014	95.90000	702.2000	94.00000	1211.000
2015	100.0000	693.8000	100.0000	1278.900
2016	108.2000	731.3000	107.0000	1453.800
2017	114.5000	809.5000	109.3000	1671.700
2018	124.6000	943.1000	112.2000	2110.800
2019	134.6000	1265.600	112.4000	2851.300
2020	143.0000	1564.300	112.9000	3537.600
2021	145.6000	1687.900	112.8000	3919.500
2022	150.8000	1689.600	115.0000	4185.600
2023	157.0000	1637.200	117.7000	4331.600
2024	169.5000	1566.800	123.3000	4615.900
2025	182.1000	1529.200	128.1000	4998.000

图 1-19 变量名与样本数据输入

(2) 组描述性统计。单击最上层标题栏 Quick, 进入 Group Statistics/Descriptive Statistics/Common sample, 如图 1-20 所示。

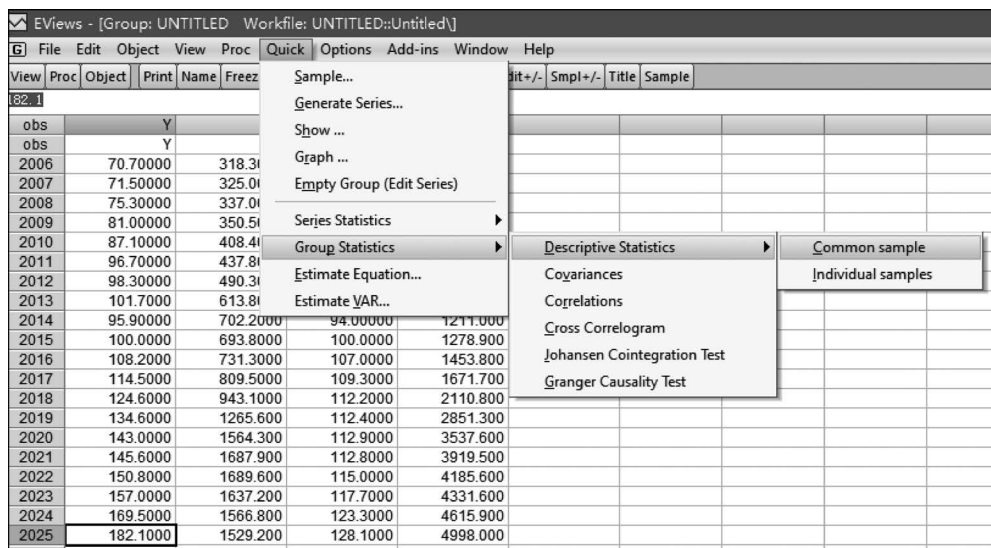


图 1-20 组描述性统计的命令

(3) 经济变量描述统计输出。出现 Series List 窗口, 输入系列名或组名, 如图 1-21 所示, 单击 OK 按钮; 得到各经济变量的描述统计量, 如图 1-22 所示, 自上至下依次为: 均值 (或期望值)、中位数、最大值、最小值、标准差、偏度、峰度、JB 统计值、正态分布的概率、变量总和、离差平方和、样本数。

(4) 经济变量描述统计列表。由软件 (英文) 格式转换成常规学术论文 (中文) 格式, 见表 1-1。

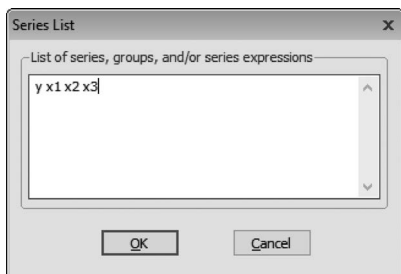


图 1-21 系列名的输入

EViews - [Group: UNTITLED Workfile: UNTITLED::Untitled]				
File Edit Object View Proc Quick Options Add-ins Window Help				
View	Proc	Object	Print Name	Freeze Sample Sheet Stats Spec
		Y	X1	X2 X3
Mean		115.4050	905.0800	113.2400 2080.970
Median		104.9500	716.7500	112.6000 1366.350
Maximum		182.1000	1689.600	146.7000 4998.000
Minimum		70.70000	318.3000	86.10000 456.8000
Std. Dev.		33.64201	528.9418	16.48905 1598.099
Skewness		0.425762	0.410712	0.263998 0.617535
Kurtosis		2.043767	1.536829	2.263764 1.793953
Jarque-Bera		1.366229	2.346339	0.684018 2.483291
Probability		0.505042	0.309385	0.710342 0.288908
Sum		2308.100	18101.60	2264.800 41619.40
Sum Sq. Dev.		21503.91	5315809.	5165.888 48524473
Observations		20	20	20 20

图 1-22 组描述性统计量的输出

表 1-1 主要变量的统计描述

变量	单位	均值	标准差	最小值	最大值
Y	亿元	115.40	33.64	70.70	182.10
X ₁	万人	905.08	528.94	318.30	1 689.60
X ₂	万元	113.24	16.48	86.10	146.70
X ₃	亿元	2 080.97	1 598.09	456.80	4 998.00

二、计量经济学模型的构建步骤

计量经济学模型构建的主要步骤如图 1-23 所示。

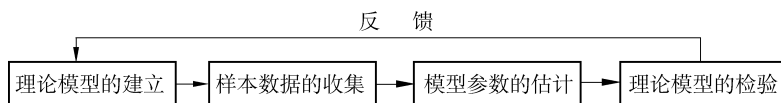


图 1-23 计量经济学模型构建的主要步骤

(一) 理论模型的建立

首先,对所研究的经济现象有足够的认识和较深入的分析;其次,确定研究目标,即明白这个计量经济学模型需要去解析哪些内在规律;再次,寻找影响研究目标的主要因素,合理确定因变量和自变量;最后,从数理统计的综合角度去构建理论模型。

例如:

$$\ln(\text{人均食品需求量}) = \alpha + \beta \ln(\text{人均收入}) + \gamma \ln(\text{食品价格}) + \delta \ln(\text{其他商品价格}) + \mu \quad (1-2)$$

式中, α 、 β 、 γ 、 δ 为参数; μ 为随机误差项。

(1) 确定模型包含的变量。根据经济学理论和经济行为分析,如式(1-2)中人均食品需求量与人均收入、食品价格、其他商品(包括替代品、互补品)价格等因素相关;考虑数据的可得性,有些变量数据很容易在国家或地方统计局网、国家与地方统计年鉴、中国知网、百度搜索、世界银行网、NBER 网等(见相关链接)找到;考虑入选变量之间的关系;要求变量间互相独立,避免变量的多重共线性。

(2) 确定模型的数学形式。利用经济学和数理经济学的成果,如生产函数、消费函数、利润函数等;根据样本数据作出的变量关系图,如菲利普斯曲线、库兹涅茨曲线、洛伦茨曲线等;选择可能的形式模拟,如一元一次方程、一元二次方程、半对数方程、全对数方程、差分方程、微分方程、指数方程等。

(3) 拟定模型中参数的期望值区间。以式(1-2)为例,预先拟定好 α 、 β 、 γ 、 δ 的正负符号、大小区间、相关关系等。

(二) 样本数据的收集

(1) 几类常用的样本数据。其包括时间序列数据、截面数据、虚变量离散数据(0,1)、面板数据等。

(2) 数据质量。其包括下列问题。

完整性问题:尽可能数据齐全,数据残缺会对模型估计结果产生误差。

准确性问题:尽可能统计资料准确,错误的数据会产生错误的结果。

可比性问题:尽可能数据统计口径统一,统计标准不同的数据放在一起比较就毫无意义,如要求汇率换算统一、货币单位统一等。

一致性问题:尽可能保证整体数据与样本数据统计的一致,尽可能保持近代数据与现代数据统计的一致等。

相关链接 1-12

部分社会经济数据网站

一、国内综合类

中国国家统计局: <http://www.stats.gov.cn/>

中国政府网: <http://www.gov.cn/shuju/>

中国统计年鉴: <http://www.stats.gov.cn/sj./ndsj/>

二、国内行业数据

中国人民银行调查统计司: <http://www.pbc.gov.cn/diaochatongjisi/116219/index.html>

中国人民银行金融市场司: <http://www.pbc.gov.cn/jinrongshichangsi/147160/147171/index.html>

中国证券监督管理委员会: <http://www.csrc.gov.cn/pub/newsite/sjtj/>

中华人民共和国文化和旅游部-统计数据: <http://zwgk.mct.gov.cn/>

中国民用航空局-统计数据: http://www.caac.gov.cn/XXGK/XXGK/index_172.html?fl=11

中国学术调查数据资料库: <http://www.cnsda.org/>

中国综合社会调查: <http://cgss.ruc.edu.cn/>

中国教育追踪调查: <https://ceps.ruc.edu.cn/>

中国宗教调查: <http://crs.ruc.edu.cn/>

中国老年社会追踪调查: <http://class.ruc.edu.cn/>

中国健康与养老追踪调查: <http://charls.pku.edu.cn/index/zh-cn.html>

中国票房数据-艺恩: <http://www.endata.com.cn/BoxOffice/>

SOSHOO 搜数-中国宏观经济数据: <http://www.soshoo.com/index.do>

中国科技资源共享网: <https://www.escience.org.cn/>

国家地球系统科学数据中心: <http://www.geodata.cn/>

中国工程科技知识中心: <http://www.ckcest.cn/home/>

三、互联网相关数据

百度指数: [http://index.baidu.com/v2/index.html# /](http://index.baidu.com/v2/index.html#/)

谷歌趋势: <https://trends.google.com/trends/explore>

巨量算数: <https://index.toutiao.com/>

360 趋势: <https://trends.so.com/?src=index.so.com>

微博指数: <https://data.weibo.com/index>

Y Combinator: <https://www.ycdb.co/>

中国互联网络信息中心: <http://www.cnnic.cn/>

四、开放数据集世界宏观数据

世界人口数据: <http://www.worldpop.org.uk>

世界人口统计: <https://www.ifreesite.com/population/>

世界银行公开数据: <https://data.worldbank.org.cn/>

世界气象组织: <https://public.wmo.int/zh-hans>

联合国粮食及农业组织-统计资料: <http://www.fao.org/statistics/zh/>

世界各地开放数据门户汇总: <https://opendatainception.io/>

OECD data: <https://data.oecd.org>

knoema 世界和地区统计数据: <https://cn.knoema.com/atlas>

全球幸福指数: <http://happyplanetindex.org/>

UN 人类发展指数: <http://hdr.undp.org/en/content/human-development-index-hdi/>

UN 儿童基金会数据中心: <http://hdr.undp.org/en/content/human-development-index-hdi/>

Statista 全球 600 多个行业、50 多个国家的商业数据: <https://www.statista.com/>

CEIC 全球宏观经济数据: <https://www.ceicdata.com/zh-hans/co>

(三) 模型参数的估计

(1) 模型参数估计方法。模型参数估计是一个纯数学技术问题,常用的有最小二乘法、二阶段最小二乘法、广义矩估计法、最大似然对数法等。

(2) 估计方法的选择。每一种模型参数估计方法都有其优点和缺点,需要根据特定的研究问题认真分析和考虑哪一种最适合,有时需要采用好几种参数估计方法进行综合分析。

(3) 应用软件。市场上有很多种可以完成计量经济学模型参数估计、模型检验、预测等基本运算的计量经济学软件包,如 EViews、SPSS、Stata、MATLAB、SAS、ET、ESP、GAUSS、MicroTSP、Data Fit、Minitab、SYSTAT、Shazam 等。目前,时间序列数据模型使用较多的软件是 EViews。

(四) 理论模型的检验

(1) 经济意义检验。以式(1-2)为例,假定变量数据的三组不同样本输入应用软件,得到

以下三种结果。

$$\ln(\text{人均食品需求量}) = -2.0 - 0.5\ln(\text{人均收入}) + 4.5\ln(\text{食品价格}) + 0.8\ln(\text{其他商品价格})$$

$$\ln(\text{人均食品需求量}) = -2.0 + 0.5\ln(\text{人均收入}) - 4.5\ln(\text{食品价格}) + 0.8\ln(\text{其他商品价格})$$

$$\ln(\text{人均食品需求量}) = -2.0 + 0.5\ln(\text{人均收入}) - 0.8\ln(\text{食品价格}) + 0.8\ln(\text{其他商品价格})$$

根据人均收入与人均食品需求量呈正比例关系,弹性系数 β 等于 $+0.5$ 是正确的;根据食品价格与人均食品需求量呈反比例关系,弹性系数 γ 是负数,但到底是 0.8 还是 4.5 呢?又根据微观经济学知识,食品价格需求弹性系数较小,通常小于 1 ,故 -0.8 是正确的。

(2) 统计检验。其由数理统计理论决定,包括拟合优度检验、总体显著性检验、变量显著性检验等,即 R^2 检验、 F 检验、 t 检验等。

(3) 计量经济学检验。其由计量经济学理论决定,包括异方差性检验、序列相关性检验、多重共线性检验、随机解释变量问题、协整检验、格兰杰因果检验等。

(4) 模型预测检验。其由模型的应用要求决定,包括稳定性检验(扩大样本重新估计)和预测性能检验(对样本外一点进行实际预测)等。

三、构建计量经济学模型的注意事项

构建计量经济学模型,需要注意三个要素:经济理论、数理方法和数据质量。

经济理论是指所研究的社会经济学问题的微观经济学理论与宏观经济学理论,是计量经济学模型构建的基础。一个不懂经济学理论、不了解经济规律的人,是不可能建立起一个哪怕是极其简单的计量经济学模型的。计量经济学家首先应该是一个经济学家。切忌在对经济问题的内部结构认识不清的情况下,想当然地设置自变量和模型的具体形式。

数理方法主要包括模型分析方法和参数计算方法,是计量经济学研究的工具,是计量经济学不同于其他经济学分支学科的主要特征。数理方法是计量经济学模型构建的重点,技术水平往往成为衡量一项研究成果等级的主要依据。人们往往过于重视数理方法的研究,而忽视对经济学理论的探讨。当然,数理方法的突破是计量经济学学科义不容辞的责任,但部分学者不了解计量经济学数理方法具体的应用背景和适用条件,陷入滥用和错用的误区。

变量样本数据反映了所研究的社会经济问题的活动水平、相互联系以及外部环境,是计量经济学研究的原料。相比之下,人们对数据质量的重视往往不够。在评审一项研究成果时,专家们往往对实证数据的可得性、可用性、可靠性关注太少。当在研究过程中出现问题时,人们较少从数据质量方面去寻找原因。目前,数据已经成为制约计量经济学实际应用的关键。

关键词

计量经济学 横截面数据 时间序列数据 面板数据 计量经济学模型

闯关习题

一、即测即练



二、简述题

1. 简述计量经济学模型,以及所包括的三个要素。
2. 简述计量经济学模型的构建步骤。

课外修炼

阅读《计量经济学导论》

《计量经济学导论》(伍德里奇著,费剑平译,第四版,中国人民大学出版社)是我国引进的国外优秀计量经济学教材之一。该书主要根据实际经验应用来解释计量经济学中的假定,并用简洁、准确的语言阐释了计量经济学研究的最新特点。

一、作者简介

伍德里奇,密歇根州立大学经济学教授,曾在国际知名期刊发表学术论文三十余篇,参与过多种书籍的写作。

二、主要特点

- (1) 不需要具备高深的数学知识,读者只要掌握部分线性代数和概率统计基础知识即可。
- (2) 强调计量经济学在实际问题中的应用。
- (3) 含有大量例题,许多是受启发于应用经济学最新有影响的作品。

有些路,可能很远,一直走下去,你会很累。可是,你想要的生活,却在很远的前方等着你!