

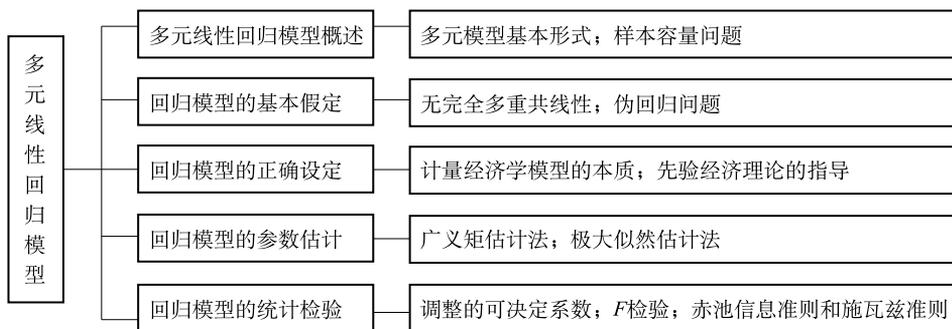
# 3

## 第三章

### Chapter 3

# 多元线性回归模型

### >>> 知识结构图



### >>> 学习目标

1. 知识目标：多元线性回归模型概述；样本容量问题；回归模型的基本假定；广义矩估计法；极大似然估计法；调整的可决定系数；方程总体线性的显著性检验等。
2. 能力目标：理解样本容量问题；理解伪回归问题；掌握广义矩估计法和极大似然估计法；区分可调整的  $R^2$  检验、 $F$  检验等。

### >>> 情景写实

在中国古代,有个国王叫齐宣王,非常喜欢听乐队吹竽,而且一定要三百人一起吹。根本不会吹竽的南郭先生,那时也请求给齐宣王吹竽。齐宣王很高兴。官府给南郭先生的待遇和那几百人一样。说明:整体吹竽效果很好,总体显著性检验( $F$ 检验)通过。

齐宣王死后,他的儿子齐湣王继承了王位。齐湣王也喜欢听吹竽,但他喜欢一个一个的独奏,南郭先生就只好灰溜溜地逃走了。说明:进行单个变量的显著性检验( $t$ 检验)时,南郭先生没有通过,被淘汰掉了。因此,在做计量实证分析时,当总体显著性检验( $F$ 检验)通过时,还有必要进行变量的显著性检验( $t$ 检验)。

一个和尚挑水喝,两个和尚抬水喝,三个和尚没水喝。这说明:变量的显著性检验( $t$ 检验)通过时,还有必要进行总体显著性检验( $F$ 检验)。

可以得出, $F$ 检验与 $t$ 检验不能相互替代,缺一不可。

## 第一节 多元线性回归模型概述

### 一、多元线性回归模型的一般形式

#### ▶ 1. 时间序列数据形式

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \mu \quad (3-1)$$

式中, $k$ 为解释变量的数目; $\beta_i$ 为模型参数; $\mu$ 为随机误差项。常数项 $\beta_0$ 可被看作一虚变量的系数,该虚变量的样本观测值始终取1,则模型中解释变量的数目为 $(k+1)$ 。

#### ▶ 2. 面板数据形式

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \mu_i$ ,改成矩阵形式为

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots \\ 1 & x_{21} & x_{22} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} \quad (3-2)$$

### 二、样本容量问题

模型参数估计是建立在被解释变量和解释变量的大量样本观测值的基础上运算完成的,其精度对样本数据的依赖性很大,通常样本数目越多精度越高。当然,也得考虑样本数据效果的有用性和时间的一致性等。例如,要分析我国当代经济发展规律,一般选择社会经济数据从1978年(党的十一届三中全会)开始。

#### ▶ 1. 最小样本容量

从普通最小二乘法原理出发,不管其质量如何,欲能得出参数估计量,所要求的样本容量的下限必须不少于模型中解释变量的数目(包括常数项),即 $n \geq k+1$ 。

#### ▶ 2. 满足基本要求的样本容量

当样本数目 $n$ 满足不少于模型中解释变量的数目(包括常数项) $k+1$ 时,但 $n$ 还是太小,除了参数估计质量不高外,一些后续统计检验工作难以实现。例如,当 $n \geq 8+k$ 时, $t$ 分布较为稳定;当 $n > 30$ 时, $Z$ 检验才能应用。

回归模型的良好性质,只有在大样本容量下才能在理论上真正得到实现。计量经济学一般实证经验表明,当 $n \geq 30$ ,或者至少 $n \geq 3(k+1)$ 时,才满足模型估计的基本统计要求。

## 第二节 回归模型的基本假定

为了使参数估计量具有良好的统计性质,特对多元线性回归模型作出若干假定。

### 一、回归模型的经典假定

**假设 1** 解释变量  $X_i$  是非随机(即固定)的,且  $X_i$  之间不存在严格线性相关(无多重共线性)。

**假设 2** 随机误差项  $\mu_i$  具有零均值、同方差及无序列相关性:

$$E(\mu_i) = 0 \quad i = 1, 2, 3, \dots, n \quad (3-3)$$

$$\text{var}(\mu_i) = \sigma_\mu^2 \quad i = 1, 2, 3, \dots, n \quad (3-4)$$

$$\text{cov}(\mu_i, \mu_j) = 0 \quad i \neq j \quad i, j = 1, 2, 3, \dots, n \quad (3-5)$$

**假设 3** 解释变量  $X_i$  与随机误差项  $\mu_i$  不相关:

$$\text{cov}(X_i, \mu_i) = 0 \quad i = 1, 2, 3, \dots, n \quad (3-6)$$

**假设 4** 随机误差项  $\mu_i$  满足正态分布:

$$\mu_i \sim N(0, \sigma_\mu^2) \quad i = 1, 2, 3, \dots, n \quad (3-7)$$

### 二、隐含的假定

**假设 5** 样本容量趋于无穷时,各解释变量的方差趋于有界常数:

$$\sum (X_i - \bar{X})^2 / n \rightarrow Q, \quad n \rightarrow \infty \quad (3-8)$$

**假设 6** 多元线性回归模型的设定是正确的。

计量经济学模型的正确设定非常重要,包括选择了正确的函数形式以及正确的自变量和因变量。

## 第三节 回归模型的正确设定

任何一项计量经济学应用研究,最重要的工作是设定计量经济学模型。只有设定了正确的总体回归模型,才能通过严格的数学过程和统计推断,得到正确的研究结果。回归模型的设定正确与否可以决定计量经济学应用研究的成败。

进入 20 世纪 70 年代,人们对计量经济学模型的预测功能提出了质疑,并不源于它未能对 1973 年和 1979 年的“石油危机”提出预报,而是几乎所有的计量经济学模型都无法预测“石油危机”对经济造成的影响。当代,翻开任何一本国际顶级经济学刊物,应用计量经济学模型分析方法的研究论文随处可见,但是人们对它的否定甚至攻击也不绝于耳。

### 一、计量经济学模型的本质

对计量经济学模型的正确设定,必须首先明确回答以下两个问题:

(1) 要确定的是经济主体内在的本质,还是经济主体之间的关系?

(2) 要确定的是经济主体之间的动力学关系,还是经济变量(经济主体经济活动的结果)之间的数据关系?

计量经济学模型分析的目的是发现和塑造整个经济世界的经济主体之间的互动规律。计量经济学模型不应该是确定在经济主体动力学关系(图 3-1)上无所指的经济变量之间的关系,经济变量及相关数据是经济主体活动的结果,脱离经济主体之间动力学关系建构的经济变量,不过是纯粹的数字。经济主体的任何行为,都应在主体和其所处的环境之间寻找原因。

构建计量经济学模型,需要注意经济理论、数理方法和数据质量三个要素。部分研究者因不了解计量模型方法的具体应用背景和适用条件,而陷入滥用和错用的误区。

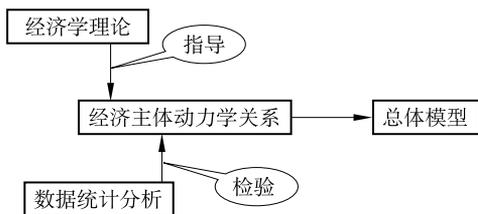


图 3-1 总体模型的构建

## 二、先验经济理论的指导

对于同一个作为研究对象的被解释变量,它和所有影响因素之间只能存在一种客观的正确关系。不同的研究者、不同的研究目的、不同的数据选择方法等,会对计量经济学模型的处理(约化和简化)不同,最终所得的应用模型自然不同,但正确的总体模型只能是一个。

### 相关链接

#### 消费者行为理论与消费函数模型

消费者行为理论是西方经济学最基本的研究。按照不同的消费者行为理论,目前存在多种消费函数模型。

##### 1. 绝对收入消费函数模型

$$C_t = \alpha + \beta_0 Y_t + \beta_1 Y_t^2 + \mu_t, \quad t = 1, 2, 3, \dots, n \quad (3-9)$$

消费  $C_t$  是由绝对收入  $Y_t$  唯一决定的,参数能较好地体现收入边际消费递减规律。

##### 2. 相对收入消费函数模型

Duesenberry 认为,在一个群体收入分布中处于低收入的个体,往往有较高的消费倾向,即

$$\frac{C_i}{Y_i} = \alpha_0 + \alpha_1 \frac{\bar{Y}_i}{Y_i} \quad (3-10)$$

由式(3-10)可以推导出

$$C_i = \alpha_0 Y_i + \alpha_1 \bar{Y}_i + \mu_i \quad (3-11)$$

Duesenberry 也认为,当前收入低于曾经达到的最高收入时,往往有较高的消费倾向,即

$$\frac{C_t}{Y_t} = \alpha_0 + \alpha_1 \frac{Y_0}{Y_t} \quad (3-12)$$

由式(3-12)可以推导出

$$C_t = \alpha_0 Y_t + \alpha_1 Y_0 + \mu_t \quad (3-13)$$

### 3. 生命周期消费函数模型

Modigliani、Brumberg 和 Ando 于 1954 年提出预算约束为

$$\sum_{t=1}^T \frac{C_t}{(1+r)^{t-1}} = \sum_{t=1}^T \frac{Y_t}{(1+r)^{t-1}} \quad (3-14)$$

当效用函数达到最大时,消费则是各个时期的收入和贴现率的函数,即

$$C_t = c_t(Y_1, Y_2, \dots, Y_T, r) \quad (3-15)$$

### 4. 持久收入消费函数模型

Friedman 于 1957 年提出持久收入消费理论,即

$$C_t = \alpha_0 + \alpha_1 Y_t^p + \alpha_2 Y_t^i + \mu_t \quad (3-16)$$

### 5. 合理预期消费函数模型

$$C_t = \alpha + \beta Y_t^e \quad (3-17)$$

式中,收入预期值  $Y_t^e$  是现期实际收入与前一期预期收入的加权和,即

$$Y_t^e = (1-\lambda)Y_t + \lambda Y_{t-1}^e \quad (3-18)$$

从上述不同的消费函数模型,可以得到如下启示:

(1) 不同的研究者,依据不同的消费理论,就可以设定不同的消费函数模型。

(2) 如果仅仅试图检验哪种消费理论适合于我国,也许是有意义的。如果研究的目的是为了揭示我国的消费行为,或者是揭示影响消费的各个因素对我国消费的实际影响,那么不同的研究者就会得到不同的结论,这样的研究是没有意义的。

(3) 先验的经济理论,可以指导我们分析实际的经济行为关系,但不能直接作为总体回归模型设定的导向。

## 三、样本数据关系的导向

计量经济学模型方法是一种经验实证的方法。一旦总体模型被设定,利用样本数据进行的数理检验,只能发现已经包含其中的哪些变量是不显著的,而不能发现未包含其中的显著变量;只能发现已经被采用的数理函数关系是不恰当的,而不能发现未被采用的正确函数关系。

总体回归模型必须反映现实的经济行为,而现实经济活动中变量之间的关系是复杂的,而且这些变量都是变化的。如果只将一部分变量引入模型,只有在其他变量不变的条件下,模型所揭示的它们与被解释变量之间的结构关系才是正确的。“其他变量不变”的条件,在现实中是无法得到满足的,所以必须将所有变量同时引入模型,因为被解释变量的变化是它们共同作用的结果。

计量经济学模型的研究范式要求对变量的设定首先做出假定,然后再进行数理检验,即各种统计检验方法。假定是从先验经济理论出发的,统计检验是从数据出发的,二者似乎在这里得到了完美的结合。但由于统计检验本质上只能证伪,不能证实,因而实际上没有被证实的假定却首先被接受了,先验的经济理论在变量设定中发挥了主导作用,再通过

统计检验的证伪也就变得毫无实际价值,自然计量经济学模型构建中的错误频发也就不足为奇了。

## 第四节 回归模型的参数估计

普通最小二乘法具有一定局限性,其参数估计值只有在回归模型满足一些严格假定条件下才具有良好的性质。广义矩估计法(Generalized Method of Moments, GMM)与极大似然估计法(Maximum Likelihood)则不需要太多严格的限制,如广义矩估计法不要求随机误差项一定非序列相关和同方差等,并且得到的回归参数估计值与真实值更接近。

### 一、广义矩估计法

广义矩估计法在大样本下是有效的,在小样本下是无效的。在随机抽样中,样本统计量将依概率收敛于某个常数,这个常数又是分布中未知参数的一个函数,即在不知道分布的情况下,利用样本矩构造方程(包含总体的未知参数),利用这些方程求得总体的未知参数。

基于模型实际参数满足一定矩条件而形成的一种参数估计方法,是矩估计方法的一般化,只要模型设定正确,则总能找到该模型实际参数满足的若干矩条件而采用广义矩估计法。

参数要满足的理论关系通常是参数函数  $f(\beta_i)$  与工具变量  $Z_i$  之间的正则条件,即

$$E[f(\beta_i)^T z_i] = 0 \quad (3-19)$$

式中,  $\beta_i$  为被估计参数。

参数估计量选择的标准是工具变量  $Z_i$  与参数函数  $f(\beta_i)$  之间的样本相关性越接近于 0 越好。用函数表示为

$$J(\beta_i) = [f(\beta_i)^T z_i]^T \mathbf{A} [f(\beta_i)^T z_i] \quad (3-20)$$

式中,  $\mathbf{A}$  为加权矩阵,任何对称正定矩阵  $\mathbf{A}$  都能得到  $\beta_i$  的一致估计。

### 相关链接

#### 广义矩估计法的参数估计

EViews 6.0 软件操作步骤:

- (1) 输入变量名和样本数据。
- (2) 选择 Quick/Estimate Equation 菜单命令。
- (3) 弹出对话框(如图 3-2 所示), Method 下拉列表中选择 GMM(广义矩估计法), Equation Specification 下的列表框中输入所有因变量和自变量名(包括常数项),在 Instrument list(工具变量列表)下的列表框中输入工具变量名。注意:工具变量的个数不能比被估参数个数少。

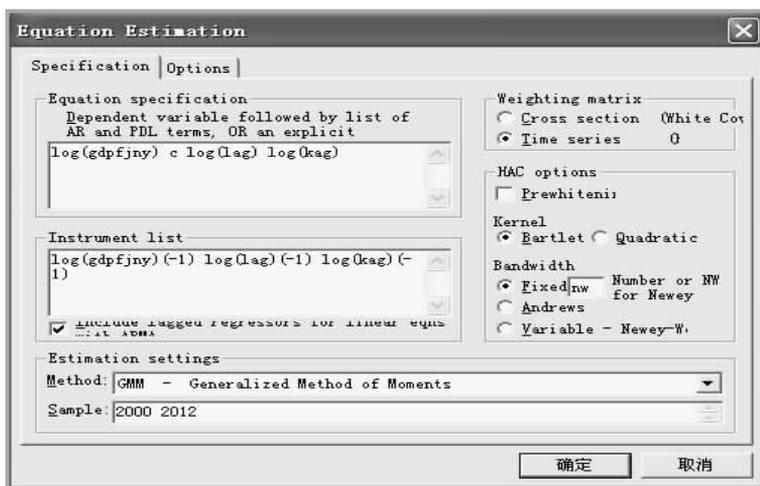


图 3-2 广义矩估计法的界面设定

在对话框右侧的 Weighting matrix(加权矩阵)选项组中有 Cross section(交叉项)和 Time series(时间序列)两个单选按钮,选择其中一个后,回归模型的异方差都是稳健的。

在 HAC options 区域的 Kernel(核函数)选项组中选择 Bartlett 或 Quadratic 单选按钮可决定计算加权矩阵时自协方差的权重,Bandwidth(宽带)选项组中有三个单选按钮,选中 Fixed(固定宽带)单选按钮后,可在其右边的文本框中输入宽带数,亦可输入“nw”,即使用 Newey-west 固定宽带。

(4) 单击“确定”按钮,输出广义矩估计法的模型参数结果(图 3-3)。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	30.82776	5.051396	6.102820	0.0001
LOG(LAG)	-3.581709	0.746948	-4.795126	0.0007
LOG(KAG)	-0.023233	0.042088	-0.552012	0.5931

R-squared	0.739667	Mean dependent var	7.382462
Adjusted R-squared	0.687601	S.D. dependent var	0.360837
S.E. of regression	0.201681	Sum squared resid	0.406754
Durbin-Watson stat	0.271097	J-statistic	0.105040

图 3-3 广义矩估计法的结果输出

## 二、极大似然估计法

极大似然估计法最早由德国数学家 Gauss 于 1821 年提出。其基本原理为:从回归模型总体随机抽取  $n$  组样本观测值后,最合理的参数估计值  $\hat{\beta}_i$  应该使得从模型中抽取该  $n$  组样本观测值的概率最大,而不是像普通最小二乘估计法那样使得模型最好地拟合样本数据。

已知某个参数估计值  $\hat{\beta}_i$  能使这个样本出现的概率最大,我们当然不会再去选择其他小概率的样本,所以干脆就把这个参数估计值  $\hat{\beta}_i$  作为参数的真实值。

求极大似然函数估计值的一般步骤:

(1) 写出似然函数:

$$P(Y_1, Y_2, \dots, Y_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \dots - \beta_i X_{ii})^2} \quad (3-21)$$

(2) 对似然函数取对数:

$$L^* = \ln p = -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \dots - \beta_i X_{ii})^2 \quad (3-22)$$

(3) 求导数。要得到  $L^*$  的极大值,则  $L^*$  分别对  $\beta_i$  求偏导数,并假定所有偏导数等于 0。

(4) 解似然方程。根据偏导数方程,求解  $\beta_i$  的估计值。

## 相关链接

### 极大似然法的参数估计

EViews 6.0 软件操作步骤:

(1) 输入变量名和样本数据。

(2) 选择 Object/New Object 菜单命令。

(3) 弹出对话框(如图 3-4 所示),Type of object 文本框中输入 LogL(极大似然估计法),在 Name for object 文本框中输入项目名(要求是字母或数字),单击 OK 按钮。

(4) 弹出如图 3-5 所示的窗口,输入描述统计语句。

以  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \mu$  为例,假定样本数  $n=30$ ,则  $n-3=27$ ,res 表示残差,var 表

示方差,又  $\hat{\sigma}_\mu^2 = \frac{\sum_{i=1}^n e_i^2}{n-3}$ ,则输入以下统计语句:

```
@logL logL1
```

```
res=Y-c(1)-c(2)*X1-c(3)*X2
```

```
var=@sum(res^2)/27
```

```
logL1=log(@dnorm(res/@sqrt(var)))-log(var)/2
```

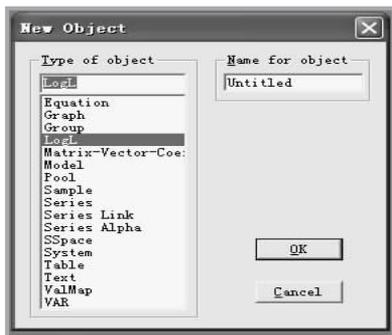


图 3-4 极大似然估计法

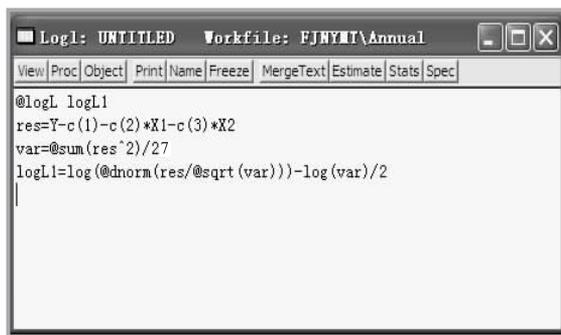


图 3-5 极大似然估计法的描述统计语句

(5) 单击图 3-5 所示窗口中的 Estimate 按钮,输出极大似然估计法的模型参数结果(图 3-6)

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-16207.92	129.9479	-124.7263	0.0000
C(2)	2386.887	13.38877	178.2753	0.0000
C(3)	154.3046	2.246226	68.69504	0.0000
Log likelihood	-65.13882	Akaike info criterion		11.35647
Avg log likelihood	-5.428235	Schwarz criterion		11.47770
Number of Coefs.	3	Hannan-Quinn criter.		11.31159

图 3-6 极大似然估计法的输出结果

## 第五节 回归模型的统计检验

### 一、调整的可决定系数

通常在回归模型中增加一个解释变量,  $R^2$  往往会增大,这是因为残差二次方和往往随着解释变量个数的增加而减少。于是给人一个错觉:要使模拟效果好,只要增加解释变量就行了。但现实情况是由于增加解释变量个数引起的  $R^2$  的增大与拟合好坏无关,因此在多元回归模型之间比较拟合优度,  $R^2$  就不是一个合适的统计检验,必须加以调整。

在样本容量一定的情况下,增加解释变量必定使得自由度减少,所以调整的思路是:将残差二次方和(RSS)与总离差二次方和(TSS)分别除以各自的自由度,以剔除变量个数对拟合优度的影响。

$$\bar{R}^2 = 1 - \frac{\text{RSS}/(n-k-1)}{\text{TSS}/(n-1)} \quad (3-23)$$

式中,  $n-k-1$  为残差二次方和的自由度;  $n-1$  为总体二次方和的自由度。

显然,如果增加的解释变量没有很好地解释能力,则对残差二次方和的减少没有多大帮助,但增加了待估参数的个数,从而使  $\bar{R}^2$  有较大幅度的下降。

在实际应用中,  $\bar{R}^2$  达到多大才算回归模型通过了检验? 没有绝对的标准,要具体情况具体分析,模型的拟合优度并不是判断模型质量的唯一标准,有时甚至为了追求模型的经济意义,可以牺牲一点拟合优度。

## 二、方程总体线性的显著性检验

前面的拟合优度检验,能够验证解释变量对被解释变量的解释程度,可以推测模型总体线性关系是否成立,但只是一个模糊的推测,不能给出一个在统计上严格的结论。这就要求进行方程总体线性的显著性检验,旨在对模型中被解释变量与解释变量之间的线性关系在总体上是否显著成立作出判断,是检验所有解释变量联合起来对被解释变量的影响。

### ▶ 1. 构造 $F$ 统计量

$$F = \frac{ESS/k}{RSS/(n-k-1)} \quad (3-24)$$

$F$  检验的思想来自总离差二次方的分解式  $TSS = ESS + RSS$ , 当  $ESS$  与  $RSS$  的比值越大时,方程总体线性关系越显著。

### ▶ 2. 提出假设

多元线性回归模型  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \mu$  中,如果方程总体上的线性关系是显著的,那么参数  $\beta_i$  显著不都为 0。于是可设原假设  $H_0: \beta_i$  全为 0; 备择假设  $H_1: \beta_i$  不全为 0。

### ▶ 3. 确定显著水平

给定一个显著性水平  $\alpha$ , 一般情况下,  $\alpha$  取 0.01(或 0.05), 即 1% 是显著的(或 5% 是显著的)。

### ▶ 4. 确定临界值

查  $F$  分布表(见附录), 得到一个临界值  $F_\alpha(k, n-k-1)$ 。

### ▶ 5. 做出判断

计算  $F$  统计量, 如果  $F >$  临界值  $F_\alpha(k, n-k-1)$ , 则拒绝原假设  $H_0: \beta_i$  全为 0, 即在  $1-\alpha$  的置信度下接受备择假设  $H_1: \beta_i$  不全为 0, 表明在  $1-1\% = 99\%$  的置信度下方程总体的线性关系是显著的。

反之,  $F \leq$  临界值  $F_\alpha(k, n-k-1)$ , 则接受原假设  $H_0: \beta_i$  全为 0, 表明在 99% 的置信度下方程总体的线性关系是不显著的。

## 三、赤池信息准则和施瓦兹准则

### ▶ 1. 赤池信息准则(Akaike Information Criterion)

$$AIC = \ln \frac{e'e}{n} + \frac{2(k+1)}{n} \quad (3-25)$$

### ▶ 2. 施瓦兹准则(Schwarz Criterion)

$$SC = \ln \frac{e'e}{n} + \frac{k}{n} \ln n \quad (3-26)$$

式中,  $e$  代表残差矩阵,  $e'$  代表残差的转置矩阵。

在被解释变量相同的条件下, 为了比较所含解释变量个数不同的多元回归模型的拟合优度, 仅当所增加的解释变量能够减少 AIC 值或 SC 值时, 在原模型中增加该解释变量才是合理的。

## 关键词

样本容量问题 广义矩估计法 极大似然估计法 调整的可决定系数 方程总体线性

的显著性检验 赤池信息准则 施瓦兹准则

## 闯关习题

### 一、单项选择题

- 按经典假设,线性回归模型中的解释变量应是非随机变量,且( )。
  - 与随机误差项不相关
  - 与残差项不相关
  - 与被解释变量不相关
  - 与回归值不相关
- 在由  $n=30$  的一组样本估计的、包含 3 个解释变量的线性回归模型中,计算得到可决定系数为 0.850 0,则调整后的可决定系数为( )。
  - 0.860 3
  - 0.838 9
  - 0.865 5
  - 0.832 7
- 下列样本模型中,哪一个模型通常是无效的( )。
  - $C_i(\text{消费})=500+0.8I_i(\text{收入})$
  - $C_i^d(\text{商品需求})=10+0.8I_i(\text{收入})+0.9P_i(\text{价格})$
  - $Q_i^s(\text{商品供给})=20+0.75P_i(\text{价格})$
  - $Y_i(\text{产出量})=0.65L_i^{0.6}(\text{劳动})K_i^{0.4}(\text{资本})$
- 模型  $\ln y_t = \ln b_0 + b_1 \ln x_t + u_t$  中,  $b_1$  的实际含义是( )。
  - $x$  关于  $y$  的弹性
  - $y$  关于  $x$  的弹性
  - $x$  关于  $y$  的边际倾向
  - $y$  关于  $x$  的边际倾向
- 调整的判定系数  $\bar{R}^2$  与多重判定系数  $R^2$  之间有如下关系( )。
  - $\bar{R}^2 = \frac{n-1}{n-k-1}R^2$
  - $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}R^2$
  - $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1+R^2)$
  - $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2)$
- 在多元线性回归模型中对样本容量的基本要求是( $k$  为解释变量个数):( )。
  - $n \geq k+1$
  - $n < k+1$
  - $n \geq 30$  或  $n \geq 3(k+1)$
  - $n \geq 30$
- 半对数模型  $\ln Y = \beta_0 + \beta_1 X + \mu$  中,参数  $\beta_1$  的含义是( )。
  - $X$  的绝对量发生一定变动时,引起因变量  $Y$  的相对变化率
  - $Y$  关于  $X$  的弹性
  - $X$  的相对变化,引起  $Y$  的期望值绝对量变化
  - $Y$  关于  $X$  的边际变化
- 下列选项中,通常所指的模型设定误差不包括( )。
  - 模型中遗漏了有关的解释变量
  - 模型中包含了无关解释变量
  - 模型中有关随机误差项的假设有误
  - 模型形式设定有误
- $F$  检验属于经济计量模型评价中的( )。
  - 统计准则
  - 经济理论准则
  - 经济计量准则
  - 识别准则
- 线性模型的影响因素( )。
  - 只能是数量因素
  - 只能是质量因素

C. 可以是数量因素,也可以是质量因素      D. 只能是随机因素

## 二、简述题

1. 多元线性回归模型满足基本要求的样本容量应该是多少?
2. 广义矩估计法的基本原理是什么?
3. 简述似然估计法的基本思路和一般步骤。
4. 简述调整的可决定系数的基本思路。
5. 简述方程总体线性的显著性检验的基本思路。
6. 简述赤池信息准则和施瓦兹准则。

## 三、软件操作题

已知某市 2002—2015 年粮食年销售量、常住人口、人均收入、肉销售量、蛋销售量、鱼虾销售量等数据如表 3-1 所示。

表 3-1 某市 2002—2015 年常住人口、人均收入及粮食等年销售量

年份	粮食年销售量 $Y$ /万吨	常住人口 $X_2$ /万人	人均收入 $X_3$ /元	肉销售量 $X_4$ /万吨	蛋销售量 $X_5$ /万吨	鱼虾销售量 $X_6$ /万吨
2002	98.45	560.20	153.20	6.53	1.23	1.89
2003	100.70	603.11	190.00	9.12	1.30	2.03
2004	102.80	668.05	240.30	8.10	1.80	2.71
2005	133.95	715.47	301.12	10.10	2.09	3.00
2006	140.13	724.27	361.00	10.93	2.39	3.29
2007	143.11	736.13	420.00	11.85	3.90	5.24
2008	146.15	748.91	491.76	12.28	5.13	6.83
2009	144.60	760.32	501.00	13.50	5.47	8.36
2010	148.94	774.92	529.20	15.29	6.09	10.07
2011	158.55	785.30	552.72	18.10	7.97	12.57
2012	169.68	795.50	771.16	19.61	10.18	15.12
2013	162.14	804.80	81180	17.22	11.79	18.25
2014	170.09	814.94	988.43	18.60	11.54	20.59
2015	178.69	828.73	1 094.65	23.53	11.68	23.37

请利用 EViews 6.0 软件,分别使用广义矩估计法和极大似然估计法构建线性回归模型  $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + \mu_t$ , 根据各解释变量的  $t$  值或  $Z$  值来剔除某些解释力较弱的变量,并重新构建回归模型。

## 课外修炼

### 阅读《计量经济学导论》

#### 一、作者简介

詹姆斯 H. 斯托克, 哈佛大学经济系教授, 加州大学伯克利分校经济学博士。曾任教于加州大学伯克利分校及哈佛大学肯尼迪政府学院。他的研究领域为经济计量方法、宏观经济预测、货币政策等, 是计量经济学领域的权威, 尤其擅长时间序列分析的研究。马克·W. 沃森, 普林斯顿大学经济系教授, 加州大学圣地亚哥分校经济学博士。他的研究领域主要包括计量经济学的时间序列分析、实证宏观经济学、宏观经济预测等。

#### 二、主要特点

(1) 将现实世界的问题和数据与理论的发展联系起来, 并认真对待实证分析中大量的重要发现。

(2) 所选取的内容反映了现代理论和实践的发展。

(3) 给出的理论和假设都与应用相符。本书的写作目的是能够指导学生在与初级课程相应的数学水平上熟练应用计量经济学, 可作为本科阶段计量经济学的入门课程来学习使用。

不管你的级别、职称和学历有多高, 教师只是学生的勤务员。