第1章

经济计量学的特征及研究范围

在经济学、金融学、管理学、营销学以及一些相关学科的研究中，定量分析用得越来越多，对于这些领域的初学者来说，掌握一至两门经济计量方面的课程是必要的——这个领域的研究变得十分流行。本章的目的旨在给初学者一个经济计量学的概貌。

1.1 什么是经济计量学

简单地说，经济计量学(Econometrics)就是经济的计量。虽然，对诸如国民生产总值(GNP)、失业、通货膨胀、进口、出口等经济概念的定量分析十分重要，但从下面的定义中，我们不难看出经济计量学的研究范围更为宽泛：

经济计量学是利用经济理论、数学、统计推断等工具对经济现象进行分析的一门社会科学。[1]

经济计量学运用数理统计知识分析经济数据，对构建于数理经济学基础之上的数学模型提供经验支持，并得出数量结果。[2]

1.2 为什么要学习经济计量学

从上述定义我们知道经济计量学涉及经济理论、数理经济学、经济统计学(即经济数据)，以及数理统计学等相关学科，但它是一门有其自己研究方向的一门独立学科。

从本质上说，经济理论所提出的命题和假说，多以定性描述为主。例如，微观经济理论中提到的：在其他条件不变的情况下(经济学中著名的 Ceteris paribus从句)，一种商品价格的上升会引起该商品需求量的减少。因而得出结论：商品的价格与该商品的需求量呈反方向变动——这就是著名的向下倾斜的需求曲线，简称需求法则。但是，该理论本身却无法度量价格和需求量这两个变量之间的数量关系，也就是说，它不能告诉我们商品的价格发生某一变动时，该商品的需求量增加或减少了多少。经济计量学家的任务就是提供这样的数量估计。换一种说法，经济计量学是依据观测和试验，对大多数经济理论给出经验的解释。如果在研究或试验中发现，当每单位商品的价格上升一美元，引起该商品需求量的下降，比如说下降 100个单位，那么，我们不仅验证了需求法则，而且还提供了价格和需求量这两个变量之间的数量估计。

---

1 Arthur S.Goldberger,*Econometric Theory*,Wiley,New York,1964，p.1.

2 P.A Samuelson,T,C.Koopmans,and J.R.N.Stone,“Report of the Evaluative Committee for *Econometrica*”,*Econometrica*, vol. 22, no. 2, April 1954,pp.141-146.

数理经济学(mathematical economics)主要关心的是用数学公式或数学模型来描述经济理论，而不考虑对经济理论的度量和经验解释。而经济计量学家感兴趣的却是对经济理论的经验确认。下面我们将会讲到，经济计量学家通常采用数理经济学家提供的数学模型，但把它们用于经验检验。经济统计学家主要关心的是收集、处理经济数据并将这些数据绘制成图表的形式。这是经济统计学家的工作：他或她收集GNP、失业、就业、价格等数据，这些数据就成为经济计量分析的原始数据。但经济统计学家却不关心用这些收集到的数据来检验经济理论。

虽然，数理统计学提供了许多分析工具，但由于经济数据独特的性质，即许多数据的生成并非可控制试验的结果，因此，经济计量学经常需要使用特殊的方法。类似于气象学，经济计量学所依据的数据不能直接控制。所以，由公共和私人机构收集的消费、收入、投资、储蓄、价格等方面的数据从本质上说是非试验性的。这就产生了数理统计学不能正常解决的一些特殊问题。而且，这些数据很可能包含了测量的误差，或是遗漏数据或是丢失数据。这就要求经济计量学家去运用特殊的方法来处理这些测量误差。

对于主修经济学和商业专业的学生来说，学习经济计量学有实用性。毕业以后，在其工作中，或许被要求去预测销售量、利息率、货币供给量或是估计商品的需求函数、供给函数以及价格弹性等等。在经济学家以专家的身份出现在联邦政府调节机构中之前，通常代表当事人或公众。而汽油和电的价格是由政府调节机构规定的，因此，这就要求经济学家能估计提议的价格的上涨对需求量(如用电量)的冲击。在这种情况下，经济学家需要建立一个关于用电量的需求函数，并根据这个需求函数估计需求的价格弹性，即，价格变动的百分比所引起需求量改变的百分比。掌握经济计量学知识对于估计这些需求函数是很有帮助的。

客观地说，在经济学和商科专业的学习与培训中，经济计量学已成为不可或缺的一部分。

1.3 经济计量学的方法论

一般说来，用经济计量方法研究经济问题可分为如下步骤：

(1) 理论或假说的陈述；
(2) 收集数据；
(3) 建立数学模型；
(4) 建立统计或经济计量模型；
(5) 经济计量模型参数的估计；
(6) 检查模型的准确性：模型的假设检验；
(7) 检验来自模型的假说；
(8) 运用模型进行预测。

为了阐明经济计量学的方法论，我们来考虑这样一个问题：经济形势会影响人们进入劳动力市场的决定吗？也就是说，经济形势是否对人们的工作意愿有影响？假设用失业率(Unemployment Rate, UNR)来度量经济形势，用劳动力参与率(Labor Forle Participation Rate, LFPR)来度量劳动力的参与，UNR和LFPR的数据由政府按时公布，那么，如何回答这个问题呢？我们按上述步骤进行分析。

1.3.1 理论或假说的陈述

首先要了解经济理论对这一问题的阐述是怎样的。在劳动经济学中，关于经济形势对人们工作意愿的影响有两个相对立的假说。一个是受挫-工人假说[discouraged-worker hypothesis (effect)]，该假说提出当经济形势恶化时，表现为较高的失业率，许多失业工人放弃寻找工作

的愿望并退出劳动市场。另一个是增加-工人假说[added-worker hypothesis (effect)]，该假说认为当经济形势恶化时，许多目前并未进入劳动市场的二手工人(比如带孩子的母亲)可能会由于养家的人失去工作而决定进入劳动市场，即使这些工作的报酬很低，只要可以弥补由于养家人失去工作而造成的收入方面的一些损失就行。

劳动力参与率的增加或减少依赖于增加工人和受挫工人的力量对比。如果增加工人的影响占主导地位，则LFPR将升高，即使是在失业率很高的情况下。相反地，如果是受挫工人的影响占主导力量，那么LFPR将会下降。我们是如何发现这一结果的呢？这只是一个实践问题。

1.3.2 收集数据

由于实验的目的，我们需要这两个变量的数量信息。一般来说，有三种统计数据可用于实践分析：

(1) 时间序列数据

(2) 横截面数据

(3) 合并数据(时间序列数据与横截面数据的联合)

1. 时间序列数据

这种数据是按时间序列排列收集得到的。比如GNP、失业、就业、货币供给、政府赤字等。数据是按照一定的时间间隔收集的 —— 每日(比如股票)，每周(比如货币供给)，每月(比如失业率)，每季度(比如GNP)，每年(比如政府预算)。这些数据可能是定量的(quantitative)(比如价格、收入、货币供给等)，也可能是定性的(qualitative)，(比如男或女，失业或就业，已婚或未婚，白人或黑人等)。我们将会发现，定性的变量(又称为虚拟变量)与定量的变量同样重要。

2. 横截面数据

横截面数据(cross-sectional data)是指一个或多个变量在某一时点上的数据的集合。例如美国人口调查局每10年进行的人口普查数据(最近的一次是在1990年4月1日)，以及密执安大学进行的夏季居民开支调查数据。这些民意调查的结果由GAllup、Harris 和其他的一些调查机构处理。

3. 合并数据

合并数据(pooled data)中既有时间序列数据又有横截面数据。例如，如果我们收集 20年间10个国家有关失业率方面的数据，那么，这个数据集合就是一个合并数据，每个国家的 20年间的失业率数据是时间序列数据，而20个不同国家每年的失业率数据又组成横截面数据。

在合并数据中有一类特殊的数据，称为 panel数据(panel data)，又称纵向数据(ongitudinal or micropanel data)。即同一个横截面单位，比如说，一个家庭或一个公司，在不同时期的调查数据。例如，美国商业局在一定时期间隔内对住房的调查。在每一时期的调查中，同样的(或居住在同一地区的)家庭被调查，以观察自上一次调查以来，其住房和经济状况是否有变化。纵向数据就是通过重复上述过程而得到的，它可对研究家庭行为的动态化提供非常有用的信息。

4. 数据来源

成功的经济计量研究需要大量高质量的数据。幸运的是国际互联网为我们提供了大量详实的数据。附录1A列出了一些网址，提供了各类微观和宏观的经济数据。学生必须熟悉这些网站并学会下载数据。当然，这些数据会不断更新，因此可得到最新的数据。

为了便于分析，这里给出一组时间序列数据。表 1-1给出了美国1980～1996年间城市劳动力参与率(Civilian Labor Force Participation Rate, CLFPR)和城市失业率(Civilian Unemployment Rate, CUNR)数据。城市失业率是指城市失业人口占城市劳动力的百分比。[1]

与物理学不同，许多收集的经济数据(比如 GNP、货币供给、道-琼斯指数、汽车销售量等)

1 我们仅考虑集合城市劳动力参与率及城市失业率，还有其他可用的数据，比如年龄，性别，种族构成等。

是非试验性的，因为数据收集机构(比如政府)或许并不直接监控这些数据。因而，劳动力参与和失业的数据来源于劳动力市场上参与者提供给政府的信息。在某种意义上，政府是数据的消极收集者。在收集数据的过程中，政府或许并不知道受挫 -工人、增加-工人假说及其他有关的一些假说。因此，收集到的数据可能是几种因素综合的结果，会影响不同个人劳动力参与的决策。也就是说，同样的数据适用于不止一个理论。

表1-1 城市劳动力参与率(CLFPR)，城市失业率(CUNR)与真实的小时平均工资(AHE82)①

年	CLFPR(%)	CUNR(%)	AHE82/$
1980	63.8	7.1	7.78
1981	63.9	7.6	7.69
1982	64.0	9.7	7.68
1983	64.0	9.6	7.79
1984	64.4	7.5	7.80
1985	64.8	7.2	7.77
1986	65.3	7.0	7.81
1987	65.6	6.2	7.73
1988	65.9	5.5	7.69
1989	66.5	5.3	7.64
1990	66.5	5.6	7.52
1991	66.2	6.8	7.45
1992	66.4	7.5	7.41
1993	66.3	6.9	7.39
1994	66.6	6.1	7.40
1995	66.6	5.6	7.40
1996	66.8	5.4	7.43

①AHE82代表1982年的平均小时工资(用美元计算)

资料来源：*Economic Report of the President，1997*,CLFPR from Table B-37,p.343,CUNR from Table B-40,p.346,and AHE82 from Table B-45,p.352.

1.3.3 建立劳动力参与的数学模型

为了观察CLFPR 与 CUNR的变动关系，首先我们作散点图(scatter diagram,or scattergram)，见图1-1。从图上可以看出，CLFPR 与CUNR呈反方向变动，(将一切情形都考虑到，或许还说明受挫工人效果比增加工人效果强)[1]。第一次估计，不妨通过这些散点做一直线并写出两者之间的简单数学模型：

$$\text{CLFPR}=B_1+B_2\text{CUNP} \tag{1-1}$$

注：Y=CLFPR　X=CUNR

式(1-1)表明城市劳动力参与率与城市失业率呈线性关系。 B_1和B_2为线性函数的参数(parameters)。[2]B_1为截距(intercept)，其值为当CUNP为零时CLFPR的值。[3]B_2为斜率(slope)，它是每一单位CUNP的变动所引起的CLFPR的变动率。更一般地，斜率度量了式右边变量每变动

1 关于这一点，参见Shelly Lungerg“The Added Worker Effect,”*Journal of Labor Economics*,vol.3,January 1985, pp.11-37.

2 一般说来，参数是一个不确定的量，可能取一组不同的值。在统计学中，通常用均值和方差等参数来描述随机变量的概率分布函数，在本书第2章中将详细地进行讨论。

3 在第5章，我们将给出回归分析中截距这一概念的一个更为准确的解释。

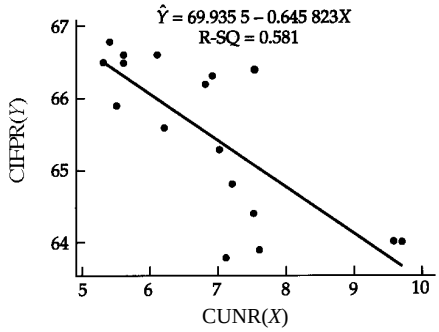


图1-1　城市劳动力参与率(百分比)与城市失业率(百分比)回归分析

一单位所引起的式左边变量的改变量。斜率值可正(若增加工人的效果大于受挫工人效果的影响)可负(若受挫工人效果的影响占主导力量)。在图 1-1中，此时的斜率为负。

1.3.4　建立劳动力参与的统计或经济计量模型

式(1-1)给出了描述城市劳动力参与率与城市失业率关系的纯数学模型，数理经济学家或许对它感兴趣，但它对经济计量学家的吸引力却是有限的。因为这样一个模型假设了变量之间的关系是精确的和确定的，也就是说，给定一个 CUNR的值，有惟一一个 CLFPR的值与之对应。在现实中，很难发现经济变量之间存在如此精确的关系。更一般地，变量之间的关系往往是不确切的或是统计的。

我们可以通过图 1-1的散点图清楚地看到这一点。虽然两变量之间存在着反方向变动关系，但两变量并非准确的或完全的线性关系，如果我们通过这 17个数据点作一条直线，并不是所有的点都准确地落在这条直线上 —— 别忘了两点确定一条直线。[1]为什么这17个数据点没有全都落在这条由数学模型所设定的直线上呢？记住这些数据是非试验性收集到的。因此，如前面所提到的那样，除了增加和受挫工人假说外，还有其他许多因素影响劳动力进入市场的决定。所以，我们所观察到的城市劳动力参与率与城市失业率之间的关系很可能是不确切的。

我们把所有其他影响劳动力参与率的因素都包括在变量 u中，于是有：

$$\text{CLFPR} = B_1 + B_2\text{CUNP} + u \tag{1-2}$$

其中，u代表随机误差项(random error term)，简称误差项(error term)。[2]u 包括了所有影响城市劳动力参与率但并未在模型中具体给出的因素(除了城市失业率)以及其他的随机因素。在本书第二部分中，我们将会看到经济计量学中的误差项与纯数理经济学的误差项有区别。

式(1-2)就是一个统计的、或经验的、或经济计量模型。更准确地说，它是一个线性回归模型(linear regression model)，这也正是本书所讨论的主题。在这个模型中，式左边的变量称为应变量(dependent variable)，式右边的变量称为自变量(independent variable)或解释变量(explanatory variable)。线性回归分析的主要目标就是解释一个变量(应变量)与其他一个或多个变量(解释变量)之间的行为关系，当然这种关系并非完全准确。

值得注意的是式(1-2)所描述的经济计量模型是来自于式(1-1)所表示的数学模型。这表明数理经济学和经济计量学是互相补充的学科。这一点可以从本书开始给出的经济计量学的定义中清楚地看到。

在继续下文之前，有一个概念值得我们注意 —— 因果关系(causation)。在式(1-2)这一回归模型中，我们说城市劳动力参与率是应变量，城市失业率是自变量或解释变量。这两个变量之

1　我们甚至试图通过这些数据点做一条抛物线，但结果与线性假定并没有什么实质性的不同。

2　在统计语言中，随机误差项有两种英文表示：random error , stochastic error

间存在因果关系(即城市失业率是原因，城市劳动力参与率是结果)吗？换句话说，回归包含因果关系吗？并不一定。正如Kendall 和Stuart 所说，“ 统计关系，无论有多强，有多紧密，也绝不能建立起因果关系：因果关系的概念必须排除在统计学之外 ”。在本例中，是根据经济理论(比如受挫工人假说)在应变量和解释变量之间来建立起原因-结果这样的关系。如果不能建立起因果关系，不妨称之为断定的关系 —— 即给定城市失业率，我们能够预测城市劳动力参与率。

1.3.5 经济计量模型参数的估计

利用表1-1给出的数据如何估计(estimates)式(1-2)中的参数B_1，B_2呢？即如何确定这些参数(parameters)的具体数值呢？这是本书第二部分的重点，在那里，我们将用适当的方法，尤其是普通最小二乘法来计算。运用最小二乘法和这些数据，得到下面的结果：

$$\widehat{CLFPR} = 69.935\ 5 - 0.645\ 8CUNR \tag{1-3}$$

注意我们在CLFPR上加一符号^提醒大家：式(1-3)是式(1-2)的估计式。图1-1是根据真实数据估计得到的回归直线。

从式(1-3)可知，B_1的估计值约为70，B_2约为－0.64。因此，平均地，如果失业率上升一个单位(比如说一个百分点)，则城市劳动力参与率将下降 0.64个百分点；也就是说，当经济状况恶化时，劳动力参与率将平均净减少 0.64个百分点，这或许表明了受挫工人效应占主导力量。我们讲“ 平均 ”是因为前面提到的误差项 u可能会导致变量之间的关系有些出入，这一点可以从图1-1清楚地看到：真实数据点并未落在估计的回归线上。回归直线上的点与真实数据点的距离称为残差。简言之，估计的回归直线(式(1-3))给出了平均城市劳动力参与率与城市失业率之间的关系——每一单位城市失业率的变化所引起的城市劳动力参与率的平均改变量是多少。常数69.93即当城市失业率为零时城市劳动力参与率的平均值。也就是说，当充分就业时(即不存在失业)，城市适龄工作人口的69.93%将参与就业。[1]

1.3.6 检查模型的准确性：模型的假设检验

式(1-3)所描述模型的准确性如何呢？通常，人们在进入劳动力市场之前，会根据一些因素，比如说失业率的大小来考虑劳动市场的状况。例如，1982年(不景气的一年)城市失业率约为9.7%，而1996年仅为5.4%。显然，失业率为5.4%，与失业率为9.7%时相比，人们更可能进入劳动力市场。但还有其他一些因素影响人们进入劳动力市场的决定。比如每小时工资或收入也是重要的决定变量。至少在短期内，工资越高越能吸引工人进入劳动力市场。其他因素也类似。为进一步说明其重要性，在表1-1中我们还给出了以美元为计量单位的真实平均小时工资(AHE82)的数据。现在加上AHE82这一影响因素，考虑下面这个模型：

$$CLFPR = B_1 + B_2CUNP + B_3AHE82 + u \tag{1-4}$$

式(1-4)是一个多元线性回归模型，而式(1-2)是一个简单的(双变量)线性回归模型。在双变量模型中，仅有一个解释变量，而在多元回归模型中有若干个(或称多元)解释变量。值得注意的是，在多元回归模型中，即式(1-4)中，同样包括误差项u，这是因为无论模型中有多少个解释变量，都不能完全解释应变量的行为。一个多元回归模型究竟需要包括多少个解释变量，须根据具体情况而定。当然，基本的经济理论通常会告诉我们哪些变量需要包括到模型之中，但是需要提醒注意的是，正如前面所提到的：回归并不意味存在因果关系；一个或多个解释变量是否与应变量存在因果关系，必须根据相关理论来判定。

如何估计式(1-4)中的参数呢？我们将在第5章和第6章讨论完双变量模型之后，在第7章中

1 然而，这仅是截据这一概念的机械的解释。我们将在第5章中阐述截据的概念。

给予详细说明。首先考虑双变量模型是因为它是多元回归模型的基础。在随后的第 7章中将会看到，多元回归模型在许多方面是双变量模型的直接延伸。

用普通最小二乘估计法估计得到回归方程(式(1-4))：

$$\widehat{CLFPR} = 97.9 - 0.446\ CUNP - 3.86AHE82 \qquad (1\text{-}5)$$

这个结果很有意思，因为两个斜率系数均为负数。负的城市失业率表明失业率每增加 1%，城市劳动力参与率将平均减少 0.44%(假设平均小时工资为一常数)。这个结果又一次支持了受挫工人假说。另一方面，若城市失业率为一常数，则平均小时工资每增加一个百分点，城市劳动力参与率将平均减少3.86%。[1] 负的平均小时工资系数有经济意义吗？为什么不期望该系数为正(即小时工资越高，则劳动力市场的吸引力也就越高)呢？我们可通过微观经济两个孪生概念，收入效应和替代效应，来验证系数为负。[2]

我们选择哪一个模型呢？式 (1-3)还是式(1-5)呢？既然式(1-5)包含(1-3)，而且增加了一个分析变量(收入)，所以我们可能选择式 (1-5)。毕竟，式(1-2)暗含地假定了除失业率以外其余变量均为常数。但是，我们的分析在哪才是尽头呢？例如，劳动力参与率可能还依赖于家庭财富，六岁以下孩子的个数(这一点对于已婚妇女考虑进入劳动市场特别重要)，孩子日托的便利程度，宗教信仰，福利事业的好坏，事业保险等等。即使这些变量的数据都可得到，我们也不会把他们都包括到模型中来，因为建模的目的不是包纳现实中的所有因素，而仅仅是一些显著因素。如果我们试图在回归模型中包括每一个可以想像到的变量，那么这个模型将会极为庞大以至于没有任何实际用处。最终选择的模型应该是对现实的合理的复制。在第 14章中我们将进一步讨论这个问题，并且探讨如何建立和发展模型。

1.3.7 检验来自模型的假设

模型最终确定之后，我们进行假设检验(hypothesis testing)。即验证估计的模型是否有经济含义，以及用模型估计的结果是否与经济理论相符。例如，受挫工人假说假设劳动力参与与失业率之间负相关。这个假说与结果相符吗？我们统计的结果与假说相一致，因为估计得到的城市失业率系数为负。

然而，假设检验或许更复杂。在这个例子中，假设得知在先前的研究中，城市失业率的系数约为-1，那么得到的结果还会与假设一致吗？如果以式 (1-3)这个模型为基础，我们可能得到一个结果，但是如果以式(1-5)模型为基础，则可能得到另一个结果。怎样解决这个问题呢？我们会在适当的章节中利用一些必要的工具来解决诸如此类的问题，但是需要提醒注意的是：根据某一特定的假说所得到的结果将依赖于最终所选择的模型。

还有一点，在回归分析中，我们不仅对模型参数的估计感兴趣，而且对检验来自于某个经济理论(或先验经验)的假设感兴趣。

1.3.8 运用模型进行预测

经过上述多个阶段，很自然地提出这样一个问题 :我们用估计的模型干什么呢？比如说式(1-5)所表示的模型。一般地，我们用模型进行预测(prediction, forecasting)。举个例子，假设现在有1997年的城市失业率和平均小时工资的数据，分别是5.2和1.2。将其带入式(1-5)，得到1997年城市劳动力参与率的预测值为 49.26%。即，如果1997年的失业率为5.2%，真实

1 式(1-5)中的城市失业率系数和平均小时工资系数称为偏回归系数 ,我们将在第7章中讨论偏回归系数的确切含义。

2 可查阅任意一本微观经济学的标准教科书。一种直接判断结果的方法是：假设夫妇双方都参加工作，则在不影响家庭收入的前提下某一方收入的大量提高，将会促使另一方撤出劳动市场。

小时工资为12美元，则该年的城市劳动力参与率约为 49%。当然，当得到1997年的城市劳动力参与率的真实值后，可与估计值做一比较，两者之间的差距代表了预测误差，从本质上说，我们希望预测误差尽可能小，这是否总是可能呢？我们将在第 6章和第7章中回答这个问题。

总结一下经济计量分析的步骤：

序　号	步　骤	例　子
1	理论的陈述	增加或受挫工人假说
2	收集数据	表1-1
3	理论的数学模型	CLFPR=B_1+B_2CUNR
4	理论的经济计量模型	CLFPR=B_1+B_2CUNR+u
5	参数估计	CLFPR=69.9－0.646CUNR
6	检查模型的准确性	CLFPR=97.9－0.446CUNR－3.86AHE82
7	假设检验	B_2<0或B_2> 0
8	预测	给定CUNR 和AHE82值，预测CLFPR 的值

虽然我们仅用劳动经济学中的一个例子来阐述经济计量学的方法论，但需指出的是我们可用同样的步骤来分析任何领域中不同变量之间的定量关系。事实上，回归分析已用于政治、国际关系、心理学、社会学、气象学和其他许多领域。

1.4 全书结构

以上我们粗略地介绍了经济计量学的特性及研究范围。本书共分四个部分。

第一部分包括第2章、第3章和第4章。主要是为那些淡忘了统计知识的读者介绍概率和统计的基础知识。本章假定读者知道一些统计学的入门知识。

第二部分向读者介绍经济计量学的基本分析工具，称之为古典的线性回归模型 (CLRM)。读者必须对古典线性回归模型有一个完整的理解，这样才能进行经济和商业领域的研究。

第三部分介绍了回归分析在实践中的运用，并讨论了当古典回归模型的假设不满足时须解决的各类问题。

第四部分讨论了联立方程模型(第 15章)。

在联立方程模型中，我们所关心的是方程中的应变量以及这些变量之间的关系。一个熟悉的例子是微观经济学中所提到的需求和供给函数。因为均衡数量和均衡价格是由需求曲线和供给曲线的交点决定的，因此我们必须同时考虑需求函数和供给函数。我们将在第15章中讨论均衡价格和均衡数量是如何确定的。

通览全书，初学者需要时刻注意，大多数论题的讨论是直接的，不牵涉数学定理和推论等。[1]学生需记住：经济计量学入门课程就像已学过的统计入门课程一样，经济计量学主要讨论的也是估计和假设检验，所不同的或者说更有意思、更有用的是被估计或检验的参数并不仅仅是均值和方差，而且还有变量之间的关系，这也正是经济学和其他社会科学所关心的。

最后一点值得提出的是，通过运用一些价格便宜的计算机软件包，会使初学者易于掌握经济计量学这门课程。读者将会在本书的学习中遇到这样的软件。一旦熟悉了一两个标准的软件，你将会意识到学习经济计量学很有意思而且会对经济学有更好的理解。

1 一些定理和推论可参见本书作者著《经济计量学基础》 (Basic Econometrics,3d ed.,McGraw-Hill,New York,1995.)

习题

1.1 假设地方政府决定在其管辖区内提高居民财产税税率。那么，将对居民住房的价格有何影响？根据本章所讨论的八个步骤来回答问题。

1.2 怎样理解经济计量学在商业和经济领域进行决策的作用？

1.3 假设你是联邦储备局主席的经济顾问，若联储主席询问你，对增加货币供给来刺激经济有何建议，那么，在你的意见中将会考虑哪些因素？在你的意见中如何运用经济计量学？

1.4 为了减少对外国石油供应的依赖，政府考虑将对汽油收取联邦税，假设福特汽车公司雇佣你分析税收增加对汽车需求量的影响，你将如何向公司提出建议？

1.5 经济计量学与纯数理经济学有何区别？它与经济理论又有何区别？

1.6 表1-2给出了美国1980～1996年间消费者价格指数(CPI)、500只股票的指数(S&P)，3月期国债利率(3-m T bill)的数据。

表1-2 消费者价格指数(CPI, 1982~1984=100),标准普尔综合指数(S&P500, 1941～1943=100)及3月期国债利率(3-m T bill, %)

年份	CPI	S&P500	3-m T bill
1980	82.4	118.78	11.506
1981	90.9	128.05	14.029
1982	96.5	119.71	10.686
1983	99.6	160.41	8.63
1984	103.0	160.46	9.58
1985	107.6	186.84	7.48
1986	109.6	236.34	5.98
1987	113.6	286.83	5.82
1988	118.3	265.79	6.69
1989	124.0	322.84	8.12
1990	130.7	334.59	7.51
1991	136.2	376.1	5.42
1992	140.3	415.74	3.45
1993	144.5	451.41	3.02
1994	148.2	460.33	4.29
1995	152.4	541.64	5.51
1996	156.9	**670.83**	**5.02**

资料来源：*Economic Report of the President*,1997,Tables B-60,p.368,B-71,p.382,and B-93,p.406.

(a) 以时间为横轴，以上述三个变量为纵轴作图。当然，你可以以3个变量分别作图。

(b) 你预计CPI 与S&P指数之间的关系如何？CPI与3月期国债利率的关系如何？为什么？

(c) 对每一个变量，根据散点图目测其回归直线。

1.7 表1-3给出了德国马克与美元之间的汇率 (一美元兑换多少德国马克)，以及两个国家的消费者价格指数。

表1-3 马克对美元的汇率及德国和美国消费者价格指数 (CPI)

年份	GM/$	美国CPI	德国CPI
1980	1.8175	82.4	86.7
1981	2.2632	90.9	92.2
1982	2.4281	96.5	97.1
1983	2.5539	99.6	100.3

(续)

年份	GM/$	美国CPI	德国CPI
1984	2.8455	103.0	102.7
1985	2.9420	107.6	104.8
1986	2.1705	109.6	104.7
1987	1.7981	113.6	104.9
1988	1.7570	118.3	106.3
1989	1.8808	124.0	109.2
1990	1.6166	130.7	112.2
1991	1.6610	136.2	116.2
1992	1.5618	140.3	120.9
1993	1.6545	144.5	125.2
1994	1.6216	148.2	128.6

资料来源： *Economic Report of the President*,1995,GM/$ from Table B-112,p.402;CPI(1982~1984=100) from Table B-110,p.400.

(a) 以时间(年)为横轴，以汇率(ER)与两个的消费者价格指数为纵轴作图。

(b) 求相对价格比率(RPR)(两国消费者价格指数之比)。

(c) 作ER与 RPR的关系图。

(d) 通过散点图作回归直线。

附录1A 万维网上的经济数据[1]

http://www.whitehouse.gov/fsbr/esbr.htm

经济统计摘要室：提供产出、收入、失业、就业、工资、生产和商业活动、价格与货币、信用和证券市场以及国际统计数据。

http://www.bog.frb.fed.us/fomc/bb/current

政府信息共享工程：提供地区经济信息； 1990年人口和住房普查； 1992年经济普查；1982，1987，1992年农业普查；1991～1995年美国进出口数据；1990年同等就业机会信息。

http://govinfo.kerr.orst.edu

国家统计局经济研究(National Bureau of Economic Research, NBER)主页：高级私人经济研究结构，提供资产价值，劳动力，生产率，货币供给，商业循环指示器等。 NBER还与其他网站有许多联系。

http://www.nber.org

panel 研究:提供一些具有代表性的个人和家庭纵向调查的数据。这些数据自从 1968年起，每年收集一次。

http://www.umich.edu/~psid

网上经济学家资料：提供有关经济活动的综合信息和数据。该网站对经济学者非常有用。

http://econwpa.wwwstl.edu/EconFAQ/EconFaq.html

联邦网景：提供联邦政府各部门几乎所有信息。

http://www.law.vill.edu/Fed-Agency/fedwebloc.html

WebEC：提供详细的经济事件与经济数据。

http://www.amex.com/

经济分析局主页：这是美国商业部门调查机构，发行杂志《商业调查》 (*Survey of Current Business*)，提供各种经济活动数据。

http://www.bea.doc.gov/

商业周期站点：提供256**组不同的经济时间序列数据。**

http://www.globalexposure.com/bci.html

CIA**出版物：提供世界事件年鉴** (World Fact Book)**和国际统计手册** (Hand book of International statistics)**。**

http://www.odic.gov/cia/publications/pubs.html

能源信息管理(DOE)**：提供各类燃料的经济信息和数据。**

http://www.eia.doe.gov/

FRED**数据库：** St.Louis **联邦储备银行公布经济和社会的历史数据，包括货币利率和商业乘数，汇率等。**

http://www.stls.frb.org/fred/fred.html

国际商业管理：提供大量的贸易统计网站以及跨国项目等。

http://www.ita.doc.gov/

STA-USA**数据库：由国家贸易数据银行提供最详细的国际贸易数据以及出口信息。另外，还提供其他国家的人口统计、政治、社会和经济状况方面的数据。**

http://www.stat-usa.gov/BEN/databases.html

网上统计资源：提供经济指数，消费者价格等数据。

http://www.lib.umich.edu/libhome/Documents.centers/stecon.html

劳动统计学：提供有关就业、失业和收入的数据以及其他有关统计的网站。

http://stats.bls.gov:80/

美国人口调查局主页：提供有关收入、就业、收入分布的数据。

http://www.census.gov/

社会调查：提供从 1972**起美国家庭个人的调查数据。有大约** 35 000**人次回答约**2 500**个不同的问题。**

http://www.icpsr.umich.edu/GSS/

贫困研究机构：一些大学中的研究中心提供的有关贫困和社会不平等调查数据。

http://www.ssc.wisc.edu/irp/

社会安全管理：这是一个官方的网站，提供各类数据。

http://www.sa.gov/

1 **摘自** Annual Editions:Microeconomics 98/99,ed.Don Cole,Dushkin/McGraw-Hill ,Connecticut,1998**。需强调的是列出的这些网址只是一部分而已。此处提供的数据在不断地更新。**

第一部分

概率与统计基础

这一部分包括3章内容。主要回顾了理解经济计量学所必备的统计理论基础知识。

第2章回顾了概率、概率密度和随机变量等基本概念。

第3章讨论了经济计量学中广泛应用的 4个重要概率分布：(1)正态分布；(2)χ^2分布；(3)t分布；(4)F分布。概括了上述4个分布的主要特征。通过几个具体例子阐明了这些概率分布是构建许多统计理论及实践的基础。

第4章介绍了古典统计学的两个重要分支：估计与假设检验。对这两个概念的正确理解对以后的学习大有帮助。

这一部分的写作风格是非正规的，但其信息量却是充实的，因为我们的目的是帮助读者温习统计学的基础知识。读者在学习经济计量学之前，可能对统计学有不同程度的了解，这3章内容将完整而自成体系地对统计学作一简要介绍。

我们将通过几个具体的例子来说明文中出现的一些基本概念。

第 2 章

基本统计概念的回顾

本章和随后两章主要是回顾一些基本的统计概念。这些概念对于理解此书是十分必要的。对于有一定统计学基础的学生来说，这 3章可作为复习课程；对于那些淡忘了统计学知识学生来说，这3章与本书剩余部分的内容一起构成一个统一的框架。建议那些统计知识较薄弱的学生，在学习的过程中阅读有关的参考书(在本章的最后给出了部分参考书目)。注意：第 2章到第4章讨论的内容并不完善，也决不是基础统计学教程。它仅仅是通向经济计量学的一座桥梁。

2.1 一些符号

我们可以用简单的数学符号表示一些数学表达式。

2.1.1 求和符号

通常用希腊字符 Σ 表示求和，其表达式为：

$$\sum_{i=1}^{i=n} X_i = X_1 + X_2 + \cdots + X_n$$

其中i为求和指数，等式的左边代表“把变量 X从第一个值(i=1)加到第 n(i=n)个值”。X_i代表变量X的第i个值。完整的求和符号为：

$$\sum_{i=1}^{i=n} X_i \left(\text{或} \sum_{i=1}^{n} X_i\right)$$

通常简单地记为：

$$\sum X_i$$

当求和的上限和下限已知或容易决定时，可表示为：

$$\sum_{X} X$$

即对所有的X的值求和。我们将会交替使用这些符号。

2.1.2 求和符号的性质

1. 若k为常数，则有：

$$\sum_{i=1}^{i=n} k = nk$$

即常数的n次求和等于该常数的n倍。因此有：

$$\sum_{i=1}^{4} 3 = 4 \times 3 = 12$$

其中，$n=4$，$k=3$。

2. 若k为常数，

$$\sum kX_i = k\sum X_i$$

即可将常数放在求和符号前。

3. $\sum (X_i + Y_i) = \sum X_i + \sum Y_i$

即对两个变量的和求和等于对两个变量分别求和的和。

4. $\sum (a + bX_i) = na + b\sum X_i$

其中a，b为常数，利用性质1、性质2、性质3可得。

我们将在本书中经常使用求和符号。

现在讨论概率论中的一些重要概念。

2.2 试验、样本空间、样本点和事件

2.2.1 试验

第一个重要概念是统计试验或随机试验(statistical or random experiment)。“随机试验是指至少有两个可能结果，但不确定哪一个结果会出现的过程。”[1]

例2.1

抛一枚硬币，掷一颗骰子和从一副纸牌中抽取一张，都是随机试验的例子。在这些随机试验中，暗含地假定了必须满足一定的条件。例如，假定硬币和骰子是正规的，没有注铅。抛一枚硬币可能出现正面朝上或正面朝下，掷一颗骰子，朝上的一面可能是1，2，3，4，5，6中的某一个。注意，试验之前并不能确定哪一个结果会出现。通过这些试验或许可以建立一条规律(比如抛一枚硬币1 000次，正面朝上有多少次？)或是检验硬币是否注铅(如果抛币100次，正面朝上70次，你会认为该枚硬币注铅了吗？)。

2.2.2 样本空间或总体

随机试验所有可能结果的集合称为总体或样本空间(population or sample space)。

例2.2

考虑这样一个试验，抛两枚同样的硬币。H代表正面朝上，T代表正面朝下。则有四种结果：HH，HT，TH，TT。其中，HH代表第一枚硬币和第二枚硬币都正面朝上，HT代表第一枚硬币正面朝上，第二枚硬币正面朝下。如此类推。

在这个例子中，全部的结果，或样本空间(总体)为 4 —— 没有其他合乎逻辑的可能的结果

1 Paul Newbold,*Statistics for Business and Economics*,4thed.,Prentic-Hall,Englewood Cliffs,N.J.,1995,p.75.

(不必担心硬币会立起来)。

例2.3

在一种双回合游戏中，O_1表示两个回合全部获胜；O_2表示第一回合获胜，第二回合失败；O_3表示第一回合失败，但第二回合获胜；O_4表示两个回合均失败。

在这里，样本空间有4种结果组成：(O_1，O_2，O_3，O_4)。

2.2.3 样本点

样本空间(或总体)的每一元素，即每一种结果称为样本点(sample point)。在例2.2中，HH，HT，TH，TT均为一样本点，同样在例2.3中，O_1，O_2，O_3，O_4也均为样本点。

2.2.4 事件

随机试验的可能结果组成的集合称为事件(events)，它是样本空间的一个子集。

例2.4

事件A表示抛两枚硬币一枚正面朝上，一枚正面朝下。从例 2.2中我们可以看到，只有HT和TH属于事件A(注：HT和TH是样本空间HH，HT，TH，TT的一个子集)。事件B表示两枚均正面朝上。很明显，只有 HH属于事件B(HH也是样本空间HH，HT，TH，TT的一个子集)。

如果两个事件不能同时发生，则两个事件称为是互斥的(mutually exclusive)。在例2.3中，如果O_1发生，即在两次游戏中均获胜，那么其他三种结果就不可能发生。如果我们确信一个事件的发生与另一事件的发生的可能性相同，则两个事件称为等可能性的(equally likely)。例如，抛一枚硬币，正面朝上和正面朝下是等可能出现的。如果可穷举试验的所有可能结果，事件称为穷举事件(collectively exhaustive)。在抛两枚硬币的例子中，因为HH，HT，TH，TT是仅有的可能结果，因此它是一个穷举事件。同样的，在例2.3中，O_1，O_2，O_3，O_4是仅有的可能结果，因此它也是一个穷举事件，当然，除非遇到下雨或自然灾害，就像1989年在旧金山举行世界锦标赛期间发生地震那样。

2.3 随机变量

虽然试验的结果可用文字来描述，比如正面朝上或正面朝下，或是黑桃 A等，但是如果将试验的结果数量化，即将试验结果和具体数字对应起来，则更为简单。在随后我们将会看到，为了统计的方便这种替代是大有益处的。

例2.5

再来看例2.2。我们不用HH，HT，TH，TT 描述试验结果，若“变量”表示了抛两枚硬币正面朝上的个数。[1]有如下情况：

第一枚硬币	第二枚硬币	正面朝上次数
T	T	0
T	H	1
T	H	1
H	T	1
H	H	2

1 一般地，变量是任意一个可变的量。更准确地说，变量是指在一个给定的集合内，可以取任一值的量。
*1英寸=0.254m

我们称变量“正面朝上个数”为一随机变量(stochastic or random variable，用符号r.v表示)。更一般地，把取值由试验结果决定的变量称为随机变量。在上例中，随机变量 r.v可取3个不同的值——0，1，2。在例2.3中，随机变量(获胜的次数)，同样可取3个不同值——0，1，2。

习惯上，通常用大写字母X，Y，Z或X_1，X_2，X_3等表示随机变量。

随机变量可能是连续的，也可能是离散的。离散型随机变量(discrete random variable)只能取到有限多个(或是可列有限多个)数值。抛两枚硬币正面朝上的次数仅能取 0，1，2，所以它是一个离散型随机变量。与此类似，获胜的次数也是一个离散型随机变量，因为它仅能取 0，1，2三个值。另一方面，连续型随机变量(continuous random variable)可以取某一区间范围内的任意值。例如，人的身高就是一个连续型随机变量，它可以取在 60～72英寸范围内的任一值。类似的，体重、降雨量、温度等都可看做是连续型随机变量。

2.4 概率

定义了试验、样本空间、样本点、事件和随机变量之后，现在我们来看另一重要概念——概率。首先我们定义事件的概率，然后扩展到随机变量的概率。

2.4.1 事件的概率：古典定义或先验定义

如果一随机试验的n个结果互斥且每个结果等可能发生，并且事件 A含有m个基本结果，则事件A的发生的概率(probability)即$P(A)$就是：

$$P(A)=\frac{m}{n}=\frac{\text{有利于事件}A\text{的基本结果的个数}}{\text{所有基本结果总数}} \tag{2-1}$$

这个定义有两个特征：

(1) 试验的结果必须互斥——即它们不能同时发生。

(2) 试验的每个结果等可能发生——例如，掷一颗骰子出现任何一个数字的机会均等。

例2.6

掷一颗骰子，有六种可能的结果：1，2，3，4，5，6。这些结果互斥，因为不可能同时出现两个或更多个数字同时朝上的结果。而且，这六种结果等可能发生。因而，根据古典概率的定义，任何一个数字朝上的概率为 1/6——因为共有六种可能结果，每种结果等可能发生，这里，m=1，n=6。

类似地，抛一枚硬币，正面朝上的概率为 1/2。因为共有两种可能的结果，H和T，而且每一种结果等可能发生。同样，在一副有 52张的扑克中，抽取任意一张的概率为 1/52。(为什么？)抽取一张是黑桃的概率为 13/52。(为什么？)

上述例子说明为什么概率的古典定义又称为先验定义(priori definition)。因为这些概率来自于纯粹的演绎推理。我们没有必要抛一枚硬币来证明正面朝上的概率为 1/2，因为它们是合乎逻辑的仅有的结果。

但是古典定义有其缺陷。如果试验的结果不是有限的或不是等可能发生的，又会怎样呢？举个例子，次年国民生产总值的概率为多少呢？或次年经济衰退的概率有多大？古典定义无法回答类似这样的问题。一个更为广泛运用的定义——概率的频率定义，[1]能够解决诸如此类的问题。

1 还有一种概率的定义，称为主观概率，它是贝叶斯统计的基础。在这种主观的或“信仰程度”概率定义下，我们可问这样的问题：伊拉克将建立民主政府的概率有多大？美国在三年一次的划艇世界锦标赛上获得冠军的概率有多大？股票市场在 2000年崩溃的概率是多少？

2.4.2 概率的频率定义或经验定义

为了介绍这个概念，我们先来看下面这个例子。

例2.7

表2-1 给出了200个学生微观经济学的考试成绩的分布，表2-1就是一个**频率分布**(frequency distribution)的例子，在这个例子中，表示了考试分数的分布。表中第3列的数字称为**频数**(absolute frequencies)，第4列的数字称为**频率**(relative frequencies)，即频数除以出现的总数(在本例中为200)。因此，分数位于70～79之间的频数为45，但频率为0.225，即用45除以200得到。

表2-1 200个学生微观经济学考试分数的分布

分 数 (1)	区间均值点 (2)	频数 (3)	频率 (4)=(3)/200
0～9	5	0	0
10～19	15	0	0
20～29	25	0	0
30～39	35	10	0.050
40～49	45	20	0.100
50～59	55	35	0.175
60～69	65	50	0.250
70～79	75	45	0.225
80～89	85	30	0.150
90～99	95	10	0.050
		总计：200	1.0

我们能把频率当作概率吗？直观地看，如果观察次数足够多，则把频率视为概率是合理的。这正是概率的经验(或频率)定义的本质所在。

更正规的，如果在n次试验(或n个观察值)中，m次有利于事件A，假定试验的次数n足够多(技术上讲，是有限的)，那么，事件A的概率$P(A)$就简单地等于m/n(即频率)，[1]需要注意的是，与概率的古典定义不同，我们无须要求试验的结果互斥，也不要求每种结果等可能发生。

简言之，如果试验的次数足够多，频率就很好地测度了(事件发生的)真实概率。因此，在表2-1中，我们把第4列中的频率看做概率。

概率的性质

事件的概率有如下一些重要性质：

(1) 事件的概率在0～1之间。因而，事件A的概率满足：

$$0 \leq P(A) \leq 1 \tag{2-2}$$

若$P(A)=0$，即事件A不会发生；若$P(A)=1$，则事件A必定发生。一般概率值介于0～1之间，如表2-1中的概率值。

(2) 若事件A，B，C，…为互斥事件，则事件和的概率等于事件概率之和，用符号表示为：

$$P(A+B+C+\cdots)=P(A)+P(B)+P(C)+\cdots \tag{2-3}$$

3. 若事件A，B，C，…为互斥事件，且为一完备事件组，则事件和的概率为1。用符号表示为：

1 究竟多少算是足够多，需要依据问题的具体情况而定。有时30次就认为相当多了。在总统竞选的民意调查中，对最终结果的预测需要800个样本(人次)就可认为是相当精确的了，虽然实际上投票的人数超过数百万。

$$P(A+B+C+\cdots)=P(A)+P(B)+P(C)+\cdots=1 \tag{2-4}$$

例2.8

在例2.6中，我们知道任一数字朝上的概率均为1/6，因为共有六种等可能发生结果。由于1，2，3，4，5，6组成一完备事件组，则$P(1+2+3+4+5+6)=1$。因为$P(1+2+3+4+5+6)=P(1)+P(2)+\cdots+P(6)=1/6+1/6+1/6+1/6+1/6+1/6=1$。

还有一些常用的性质：

(1) 事件A，B，C，…称为相互独立的事件，如果事件积的概率等于事件概率的积，用符号表示：

$$P(ABC\cdots)=P(A)P(B)P(C)\cdots \tag{2-5}$$

其中，$P(ABC\cdots)$表示事件A，B，C，…同时发生或联合发生，因此，$P(ABC\cdots)$称为联合概率(joint probability)。与联合概率相对应，$P(A)$，$P(B)$，$P(C)$，…称为非条件概率(unconditional)，或边缘概率(marginal)或单独概率(individual)。在2-6节中将详细讨论。

例2.9

假设同时抛两枚硬币。那么两枚均正面朝上的概率是多少？令事件A表示第一枚正面朝上，事件B表示第二枚正面朝上，因此现在要求概率$P(AB)$。一般地认为第一枚正面朝上的概率独立于第二枚正面朝上的概率，因而有，$P(AB)=P(A)P(B)=(1/2)(1/2)=1/4$，这里抛币一次正面朝上的概率为1/2。

(2) 如果事件A，B，C，…不是互斥事件，则式(2-5)需要重新定义。若事件A，B不是互斥事件，则有：

$$P(A+B)=P(A)+P(B)-P(AB) \tag{2-6}$$

其中，$P(AB)$为事件A，B同时发生的联合概率。当然，如果A，B互斥，则$P(AB)=0$(为什么？)即为式(2-3)。式(2-6)很容易推广到两个以上事件的情况。

例2.10

从一副扑克中抽取一张，则是红心或是皇后的概率是多少？很显然，抽红心和抽皇后不是互斥事件，因为4张皇后中有一张是红心。因而有：

$$\begin{aligned}P(\text{或是红心或是皇后})&=P(\text{红心})+P(\text{皇后})-P(\text{既是红心又是皇后})\\&=13/52+4/52-1/52\\&=4/13\end{aligned}$$

若有事件A，B。现在我们想求在事件B发生的情况下，事件A发生的概率。这种概率称为在事件B发生条件下事件A的条件概率(conditional probability)。用符号$P(A\mid B)$表示，我们用下面的公式来计算：

$$P(A\mid B)=\frac{P(AB)}{P(B)} \tag{2-7}$$

即给定事件B事件A发生的条件概率等于事件A、B的联合概率与事件B的边缘概率之比。即，

$$P(B\mid A)=\frac{P(AB)}{P(A)} \tag{2-8}$$

例2.11

会计入门班有500个学生，其中男生300人，女生200人。在这些学生中，100个男生和60个女生计划主修会计学。现在，随机抽取一人，发现这个学生计划主修会计学。那么，这个学生是男生的概率是多少？

令事件A代表学生是男生，事件B代表主修会计学的学生。因此，我们要求概率$P(A \mid B)$。从条件概率的公式得：

$$P(A \mid B)=\frac{P(AB)}{P(B)}$$
$$=\frac{(100/500)}{(160/500)}$$
$$=0.625$$

从给出的数据我们很容易得到$P(A)=300/500=0.6$，即抽取一人是男生的非条件概率为0.6，显然与上面求得的0.625不同。

这个例子告诉我们一个非常重要的结论：一般条件概率不等于非条件概率。[1]

2.4.3 随机变量的概率

我们给出了样本结果或样本空间中事件的概率，同样也可给出随机变量的概率。因为随机变量只不过是样本空间基本结果的数字表示，比如例2.5。在本书中，我们关注的主要是随机变量，比如GNP、货币供给、价格、工资等，因此，需要知道如何给出随机变量的概率。下面，我们就来讨论随机变量的概率分布。

2.5 随机变量与概率密度函数

根据随机变量X的概率分布函数或概率密度函数(PDF)，可以知道随机变量的取值及与之相对应的概率。为了便于理解，我们首先看离散型随机变量的概率密度函数，然后再考虑连续型随机变量的情形。

2.5.1 离散型随机变量的概率密度函数

前面讲过，离散型随机变量仅可取有限个(或有限可列个)数值。为了理解离散型随机变量的概率密度函数(PDF)，再看例2.5。

例2.12

随机变量X代表抛两枚硬币正面朝上的次数，现考虑下表：

正面朝上的次数 X	PDF $f(X)$
0	1/4
1	1/2
2	1/4
	———
	1.00

在这个例子中，随机变量X取3个不同值0，1，2。X取0值的概率为1/4(即抛两枚硬币没有一次正面朝上)，因为共有4种可能结果，只有1个有利于结果TT。同样的，在四种可能结果中，只有一个有利于“两枚均正面朝上”，因而，其概率也为1/4。另一方面，两种结果HT、TH有利于事件“有一枚正面朝上”，因而，其概率为2/4=1/2。注意：这里我们是用概率的古典定义给出这些概率值。

1 但是需要指出的是：如果事件A,B相互独立，即$P(AB)=P(A)\cdot P(B)$，则$P(A|B)=\frac{P(AB)}{P(B)}=P(A)\cdot P(B)/P(B)=P(A)$。也就是说，如果两事件是相互独立的，则事件$A$在给定条件$B$下的条件概率等于其非条件概率。

上表给出了变量 X 的可能值以及与之相对应的概率值。我们用函数 $f(X)$ 表示概率分布或概率密度函数(probability distribution or probability density function，PDF)——它给出了变量 X 取不同值时的概率。由于在上面这个例子中，变量 X 是离散型随机变量，所以，上表中的 PDF 称为离散型随机变量的概率密度函数(PDF of a discrete random variable)。附带地提醒一句，上表中的概率和为1，因为这三种结果完全包括了所有的可能情况。(回顾2.4节性质3)。

更正规的，函数

$$f(X)\begin{cases} = P(X = x_i) & \text{当 } i = 1,2,3\cdots,n \\ = 0 & \text{当 } X \neq x_i \end{cases} \tag{2-9}$$

称为离散型随机变量的概率密度函数。其中，$P(X = x_i)$ 表示离散型随机变量 X 取 x_i 时的概率值。因此，在上例中，$P(X=2)$ 表示随机变量“正面朝上的次数”为 2 时的概率。图2-1给出离散型概率密度函数的几何图形。

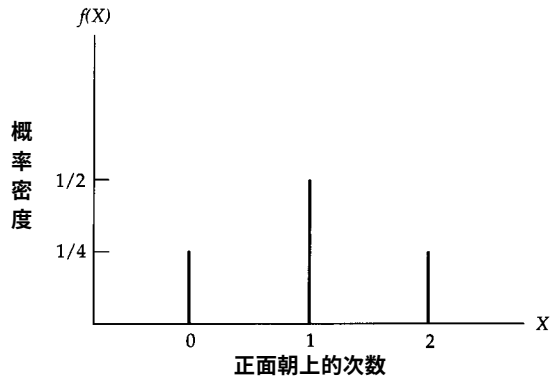


图2-1　“抛两枚硬币正面朝上”的概率密度函数 (例2.12)

2.5.2　连续型随机变量的概率密度函数

连续型随机变量的概率密度函数(PDF of a continuous random variable)的概念与离散型随机变量的相类似，所不同的是，现在我们度量的是随机变量在某一特定范围或区间内的概率。令 X 代表一连续型随机变量——身高，用英寸来度量，假设我们欲求“人的身高”在某一区间内(比如说60～68英寸)的概率，进一步假定随机变量“身高”的概率密度函数见图 2-2。[1]

图2-2中的阴影部分给出了身高在60～68英寸的概率。(实践中，我们如何度量这个概率呢？在第3章中给予讨论)。有一点需要指出：连续型随机变量取某一特定值(比如取值 63英寸)的概率为0。我们总是在一个区间内度量连续型随机变量的概率，比如说，在区间 62.5～63.5 英寸内。

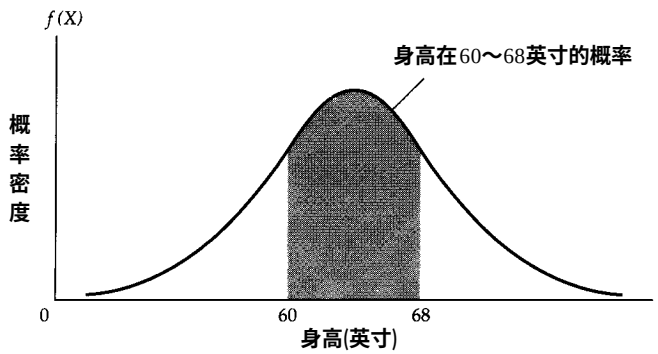


图2-2　连续型随机变量的概率密度函数

2.5.3　累积分布函数

与随机变量 X 的概率密度函数相对应，$F(X)$ 称为累积分布函数(cumulative distribution

1　在第3章中将会看到，图示的概率密度函数是著名的正态概率分布。

function，CDF)，定义如下：

$$F(X)=P(X\leq X) \tag{2-10}$$

其中，$P(X\leq X)$表示随机变量X取小于或等于x的概率。(当然，对于连续型随机变量取某一特定值的概率为0)。因此，$P(X\leq 2)$表示变量X 取小于或等于2的概率。

例2.13

抛币4次，求随机变量(正面朝上的次数)的概率密度函数和累积分布函数。

正面朝上的次数(X)	PDF		CDF	
	X值	$f(x)$ (PDF)	X值	$f(x)$ (CDF)
0	0≤X<1	1/16	X≤0	1/16
1	1≤X<2	4/16	X≤1	5/16
2	2≤X<3	6/16	X≤2	11/16
3	3≤X<4	4/16	X≤3	15/16
4	4≤X<5	1/16	X≤4	1

根据累积分布函数的定义，累积分布函数仅仅是当 X的值小于或等于某一给定 x 时的概率密度函数的“累积”或简单求和。即，

$$F(X)=\sum^{x} f(X) \tag{2-11}$$

其中，$\sum^{x} f(X)$表示对X的值小于或等于给定的x的所有概率密度函数求和。因此，在本例中，X取小于2的概率为5/16，但X 取小于3的概率为11/16。当然，X取小于4的概率为1。(为什么？)

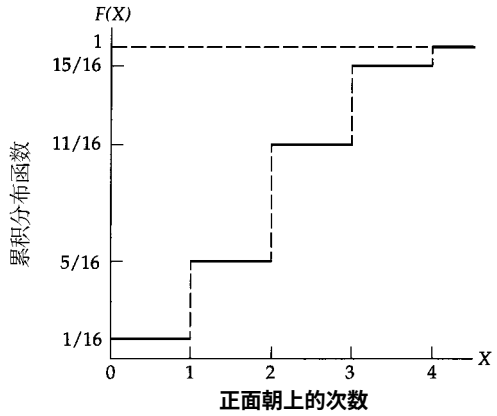


图2-3 离散型随机变量的累积分布函数 (例2.13)

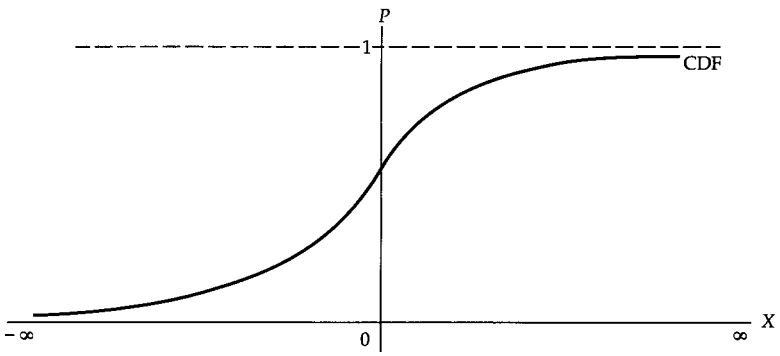


图2-4 连续型随机变量的累积分布函数

图2-3给出了例2.13的累积分布函数。因为例2.13中的随机变量是离散型的，故其累积分布函数也是非连续的，称之为分段函数(step function)。如果随机变量是连续型的，则其累积分布函数是一条连续的曲线，如图2-4所示。

2.6 多元随机变量的概率密度函数

到目前为止，我们一直关注的是单变量或一元随机变量的概率密度函数。例2.12和2.13都是单变量概率密度函数，因为在那里仅仅考虑一个随机变量，比如说“抛两枚硬币正面朝上的次数”或“抛币四次正面朝上的次数”。但是我们不必仅限于此，我们可以用不止一个的随机变量来描述一个试验的结果，在此情况下，求得的概率密度称为多元(多维)概率密度(multivariable probability distributions)。最简单的多元概率密度函数是双变量概率密度函数。我们用一个具体例子来说明。

例2.14

表2-2给出了50支债券的债券等级(X)及收益率(Y)数据，其中X有三个不同水平：X=1(Bbb)，X=2(Bb)，X=3(B)。根据标准普尔债券等级评定，Bbb,Bb,B都是中等信用的债券；Bb的信用略高于B，而Bbb又略高于Bb。即字母越少，股票的风险越大。

表2-2 双变量的频数分布：债券等级(X)与债券收益(Y)

等级(X) / 收益(Y)(%)	1 (Bbb)	2 (Bb)	3 (B)	总计
8.5	13	5	0	18
11.5	2	14	2	18
17.5	0	1	13	14
合计	15	20	15	50

表2-3 双变量概率密度

等级(X) / 收益(Y)(%)	1 (Bbb)	2 (Bb)	3 (B)	总计
8.5	0.26	0.10	0.00	0.36
11.5	0.04	0.28	0.04	0.36
17.5	0.00	0.02	0.26	0.28
合计	0.30	0.40	0.30	1.00

在这个例子中，有两个随机变量X(债券等级)和Y(债券收益)。从上表可知有13种债券等级为1(也即3个B)，其收益率为8.5%。与此类似，有14种债券等级为2(也即2个B)，其收益率为11.5%，等等。换句话说，表2-2给出了两个变量X和Y的频数分布。

我们把表2-2中的每一个数值都除以50，将频数转化为相对频率，即概率(为了简单起见，假设总体或样本空间仅由50种债券组成)。其结果见表2-3。

表2-3提供了一个双变量或联合概率密度函数(bivariate or joint probability density function, or joint PDF)。表中每一值均为联合概率(joint probability)——即变量X取一给定值(例如取2)与变量Y取一给定值(例如取11.5%)时的概率。通常用$f(X, Y)$表示联合概率密度函数。因为在本例中X，Y仅取离散值，故这是一个离散型联合概率密度函数。

更正规的，令X、Y是两个离散型随机变量，那么函数：

$$f(X, Y) = P(X=x, Y=y) \qquad (2\text{-}12)$$
$$= 0 \quad \text{当} X \neq x, Y \neq y \text{ 时}$$

称为离散型联合概率密度函数。它给出了当 X取x，且Y取y的概率，其中x、 y为给定的某一具体值。因此，在上例中，当 X取3且Y取17.5%时的(联合)概率为0.26。用同样的方法可以得知其他的联合概率。

两个连续型随机变量的联合概率可类似地定义，只是数学表达式比较复杂，有兴趣的读者可阅读有关的参考书。

2.6.1 边缘概率密度函数

我们已经学习了单变量的概率密度函数，例如 $f(X)$，$f(Y)$，以及双变量的或联合概率密度函数，$f(X, Y)$。它们之间有联系吗？是的，的确有。

与联合概率密度函数$f(X, Y)$相对应，$f(X)$，$f(Y)$称为单变量，非条件或边缘概率密度函数(univariate, unconditional, individual, or marginal PDFs)。当X 取一给定值(例如取2)而无论Y 取值如何时的概率称为X的边缘概率，其概率密度称为 X的边缘概率密度函数。如何计算边缘概率密度函数呢？很容易。我们看表2-3的总计这一行，X取1的概率为0.30，不管Y取值如何；X取2的概率为0.40，不管Y取值如何；X取3的概率为0.30，不管Y值如何。因此，X的边缘概率见表2-4。同样，也给出了Y 的边缘概率。需要指出的是：$f(X)$[或$f(Y)$]的概率密度函数之和为1。(为什么？)

表2-4 X与Y的边缘概率分布

X(债券等级)	$f(X)$	Y(债券收益%)	$f(Y)$
1	0.30	8.5	0.36
2	0.40	11.5	0.36
3	0.30	17.5	0.28
总计	1.00	总计	1.00

读者很容易注意到：求X的边缘概率仅需将X相对应的联合概率相加。即把表 2-3中的各列相加。同样的方法，把表 2-3中的各行相加我们可以得到 Y的边缘概率。一旦计算出边缘概率，那么可以直接地求出边缘概率密度函数。

2.6.2 条件概率密度函数

继续我们的债券等级－收益一例，现在假设我们想知道在债券等级为 1的条件下，收益为8.5%的概率是多少？换句话说，在给定 X=1的条件下， Y=8.5%的概率是多少？这就是我们所说的条件概率(conditional probability)(回忆一下我们先前有关事件的条件概率的讨论)。我们可从下面定义的条件概率密度函数(conditional probability density function)中得此概率。

$$f(Y \mid X) = P(Y = y \mid X = x) \tag{2-13}$$

其中，$f(Y \mid X)$代表Y的条件概率密度函数；它给出了在给定 X=x(例如取1)条件下 Y取y 值(例如，8.5%)的概率。类似地，给出X的条件概率密度函数：

$$f(X \mid Y) = P(X = x \mid Y = y) \tag{2-14}$$

注意：上面的两个条件概率密度函数都是对离散型随机变量而言的。因此，称之为离散型条件概率密度函数。连续型条件概率密度函数可类似定义，只不过其数学表达式有些复杂。

下面讨论计算条件概率密度函数的一种简单方法：

$$f(X \mid Y) = \frac{f(X,Y)}{f(Y)} = \frac{X\text{与}Y\text{的联合概率}}{Y\text{的边缘概率}} \tag{2-15}$$

$$f(Y \mid X) = \frac{f(X,Y)}{f(X)} = \frac{X\text{与}Y\text{的联合概率}}{X\text{的边缘概率}} \tag{2-16}$$

用语言描述就是，在给定另一变量取值条件下某变量的条件概率密度函数等于这两个变量的联合概率与另一变量的边缘或非条件概率密度函数之比。［与在事件B发生条件下事件A的条件概率，即$P(X \mid Y)$，相比较。］

来看我们的这个例子，现欲求 $f(Y=8.5 \mid X=1)$,则有：

$$f(Y=8.5 \mid X=1) = \frac{f(Y=8.5, X=1)}{f(X=1)}$$
$$= 0.26/0.30 \quad \text{(从表2-3得知)}$$
$$\approx 0.8667$$

从表2-3中可知Y 取8.5%的非条件概率是0.36，但在债券等级为1的条件下，Y 取8.5%的概率增加到0.87(近似值)。

在第5章中将会看到，在回归分析中，我们关注的是研究一个变量在给定另一个(或更多个)变量取值条件下的行为。因此，条件概率密度函数的知识对于建立回归分析非常重要。

2.6.3 统计独立性

在回归分析的研究中，另一个非常重要的概念是独立随机变量(independent random variables)，它与前面讨论过的事件的独立性有关。我们用一个具体的例子解释这个概念。

例2.15

一个袋子中放着分别写有1，2，3的三个球。现从袋子中有放回地随机抽取两球(即每次抽取一个，然后放回再抽取一个)。令变量X表示第一次抽取球的数字，Y代表第二次抽取球的数字。表2-5给出这两个变量的联合概率密度函数和边缘概率密度函数。

现考虑概率$f(X=1, Y=1)$，$f(X=1)$，$f(Y=1)$。由表2-5得知其概率值分别为1/9，1/3和1/3。其中，第一个为联合概率，而剩下的两个为边缘概率。但是，此例中的联合概率等于两个边缘概率之积。在这种情况下，我们称这两个变量具有统计独立性(statistically independence)。更一般地，两个变量X和Y称为统计独立的，当且仅当它们的联合概率密度函数可以表示成为其边缘概率密度函数之积。用符号表示为：

$$f(X, Y) = f(X) f(Y) \tag{2-17}$$

读者容易验证：在表2-5给出的其他任意X和Y值的组合，其联合概率密度函数都等于各自的边缘概率密度函数之积；即这两个变量是统计独立的。需要牢记的是：对于所有的 X和Y值组合式(2-17)均成立。

表2-5 两随机变量的统计独立性

		X			$f(Y)$
		1	2	3	
Y	1	1/9	1/9	1/9	3/9
	2	1/9	1/9	1/9	3/9
	3	1/9	1/9	1/9	3/9
	$f(X)$	3/9	3/9	3/9	1

例2.16

在例2.14中的债券等级与债券收益是独立随机变量吗？让我们用式(2-17)给出的独立变量的定义来检验。令X=1(Bbb)，Y=8.5%。从表2-3可知，$f(X=1, Y=8.5)=0.26$; $f(X=1)=0.30$，$f(Y=8.5)=0.36$。显然，在此情况下，$0.26 \neq (0.30)\times(0.36)$。因此，债券等级与债券收益不是独立的随机变量，这也并不让人感到惊奇。(为什么？)

2.7 概率密度的特征

虽然概率密度函数给出了随机变量的取值及其相应概率，但是通常我们并不对整个概率密度函数感兴趣。因此，在例2.12中，我们或许不想知道“没有正面朝上”,“一次正面朝上”或是“两次正面朝上”的概率，我们关注的是抛币若干次“正面朝上”的平均次数。换句话说，我们感兴趣的是一些综合指标，更专业地说，是概率密度的矩(moments)。两个最常用的综合指标(或矩)是期望值(概率密度的一阶矩)和方差(概率分布的二阶矩)。

2.7.1 期望值：集中趋势的度量

离散型随机变量的期望值(expected value)用符号$E(X)$表示(读作X的期望值)，其定义为：

$$E(X) = \sum_X Xf(X) \tag{2-18}$$

其中$f(X)$是X的概率密度函数，$\sum_X$表示对所有X求和。[1]

用文字描述即为随机变量的期望值是其各可能取值的加权平均，与各可能取值对应的概率为权重。或者说，随机变量的期望值就是该变量的可能取值与其概率之积的累加。随机变量的期望值也称为均值，更准确地应称为总体均值(population mean value)。(随后会讨论其中原因。)

例2.17

掷一个骰子若干次。求每个数字出现的期望值？参见前面讨论过的例2.6。结果见表2-6。

运用式(2-18)的期望值定义求出期望值为3.5。这样的结果奇怪吗？因为这是一个离散型随机变量，仅能在1，2，3，4，5，6中取一个值。在本例中，期望值或均值为3.5，表示如果掷骰子若干次，平均而言，得到的数为3.5(介于3～4之间)。

图2-5表示了上例中期望值。

表2-6 随机变量(正面朝上数字)的期望值

正面朝上的数字(1) X	概率(2) $f(X)$	数字×概率(3) $Xf(X)$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6
		$E(X)$=21/6=3.5

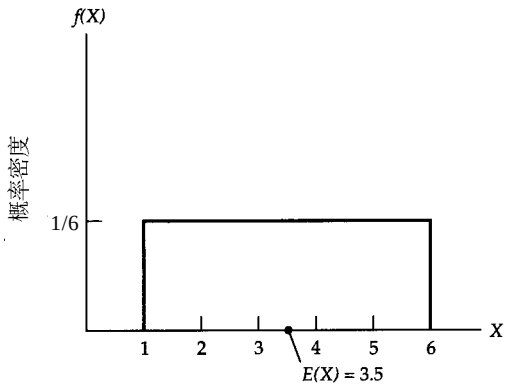


图2-5 离散型随机变量(例2.17)的期望值，$E(X)$

1 连续型随机变量的期望可类似定义，所不同的是求和符号“∑”换成积分符号“∫”。

例2.18

在债券等级一例中，债券等级的期望值是多少？我们很容易从表 2-4中得到。把变量X的各可能值与其相对应的概率之积相累加即得债券等级的期望值。

$$1(0.30)+2(0.40)+3(0.30)=2.0$$

因此，债券等级的期望值为2，恰为Bb值。

例2.19

在前一例中，债券收益的期望值是多少？我们仍旧可以从表 2-4中得到。把变量Y(收益)的各可能值与其相对应的概率之积相累加即得债券收益的期望值。

$$8.5(0.36)+11.5(0.36)+17.5(0.28)=12.10$$

这就是债券收益的期望。

期望的性质

下面给出期望值的一些重要性质，在后面的学习中，你会发现这些性质非常有用：

(1) 常数的期望值是其自身。若b为一常数，则有：

$$E(b)=b \tag{2-19}$$

例如，若b=2, 则$E(2)$=2。

(2) 两随机变量和的期望值等于两变量期望值之和。[8]给定随机变量X和Y，有，

$$E(X+Y)=E(X)+E(Y) \tag{2-20}$$

(3) 但是，$E(X/Y)\neq\frac{E(X)}{E(Y)}$ (2-21)

即两随机变量之比的期望值不等于该两个变量的期望值之比。

(4) 同样，一般情况，

$$E(XY)\neq E(X)E(Y) \tag{2-22}$$

即一般而言，两随机变量积的期望值不等于两变量期望之积。但是，有例外的情况，如果随机变量X和Y相互独立，则有：

$$E(XY)=E(X)E(Y) \tag{2-23}$$

回想一下变量X与Y相互独立的概念：当且仅当$f(X,Y)=f(X)f(Y)$时X与Y称为相互独立的。也就是说，两个变量的联合概率密度函数等于各变量的概率密度函数的乘积。

(5) 若a为常数，则有：

$$E(aX)=aE(X) \tag{2-24}$$

即随机变量的常数倍的期望值等于该变量期望的常数倍。

(6) 若a,b为常数，那么，

$$\begin{aligned}E(aX+b)&=aE(X)+E(b)\\&=aE(X)+b\end{aligned} \tag{2-25}$$

我们利用性质(1)，(2)和(5)可得到。

例如，$E(4X+7)=4E(X)+7$

2.7.2 方差：离散程度的度量

期望值简单地给出了随机变量密度的中心，但却没有表明单个值在均值附近是如何分散或分布的。最常用的度量这种分散程度的工具是方差(variance)。下面我们给出方差的定义。

1 这条性质可以推广到两个以上的变量。因此有，

$$E(X+Y+Z+W)=E(X)+E(Y)+E(Z)=E(W)$$

随机变量X的期望值为$E(X)$，为了简便，用符号u_x表示期望值(u为希腊字母)。则方差定义为：

$$\text{var}(X)=\sigma_x^2=E(X-u_x)^2 \tag{2-26}$$

其中希腊字母σ_x^2是常用的方差符号。式(2-26)表明随机变量的方差等于该变量与其均值之差的平方的期望值。因而，方差表明了随机变量 X 的各取值与其期望值或均值的偏离程度。如果所有X的值恰好都等于$E(X)$，则方差为零，但如果X的值偏离均值幅度很大，则方差也相对较大。如图2-6。注意方差不能为负。(为什么？)

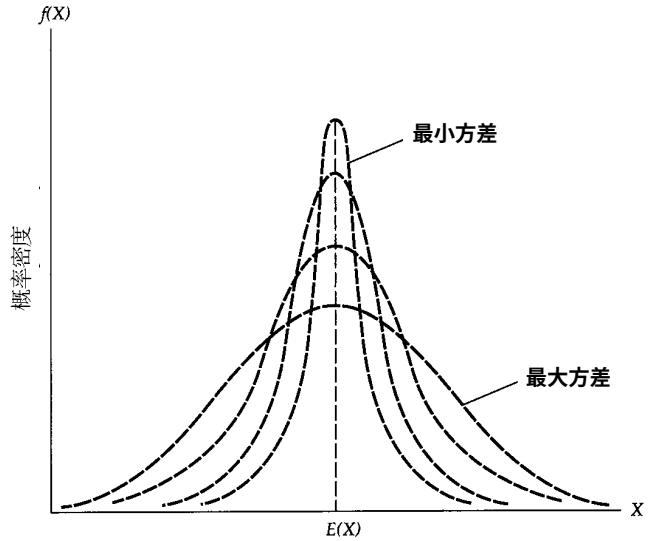


图2-6 同期望值的连续型随机变量的概率密度函数

σ_x^2的正的方根σ_x称为标准差(standard deviation,s.d).

式(2-26)给出了方差的定义式。通常我们用下式计算方差：若 X 是离散型随机变量，

$$\text{var}(X)=\sum_X (X-u_x)^2 f(X) \tag{2-27}$$

可用类似的公式计算连续型随机变量的方差。

式(2-27)表明，对于离散型随机变量，计算其方差，需先求出变量X与其期望值之差，然后求其平方，再乘以其相应概率，对变量X的每一取值重复上述过程，最后把所有求得的值相加即得。

例2.20

仍以例2.17为例，在那里我们求得变量(重复掷一颗骰子得到的正面朝上的数字)的期望值为3.5，现求其方差。我们建立表2-7。

表2-7 随机变量 X (正面朝上的数字)的方差

正面朝上的数字 X	概率 $f(X)$	$(X-u_x)^2 f(X)$
1	1/6	$(1-3.5)^2(1/6)$
2	1/6	$(2-3.5)^2(1/6)$
3	1/6	$(3-3.5)^2(1/6)$
4	1/6	$(4-3.5)^2(1/6)$
5	1/6	$(5-3.5)^2(1/6)$
6	1/6	$(6-3.5)^2(1/6)$
		求和=2.916 7

因此，其方差为2.916 7。取正的平方根，得其标准差，约为 1.707 8。

方差的性质：

(1) 常数的方差为零。根据定义，一个常数没有变异性。

(2) 如果 X 与Y是两个相互独立的随机变量，那么，

$$\text{var}(X+Y)=\text{var}(X)+\text{var}(Y) \tag{2-28}$$
$$\text{var}(X-Y)=\text{var}(X)+\text{var}(Y)$$

即两独立随机变量的和或差的方差等于两变量方差的和。

(3) 如果b是一常数，那么，

$$\text{var}(X+b)=\text{var}(X) \tag{2-29}$$

即将变量加上一个常数不改变该变量的方差。例如， $\text{var}(X+7)=\text{var}(X)$。

(4) 如果a为一常数，那么，

$$\text{var}(aX)=a^2\text{var}(X) \tag{2-30}$$

即随机变量常数倍的方差等于该变量方差的常数平方倍。例如， $\text{var}(5X)=25\,\text{var}(X)$。

(5) 如果a, b为常数，那么，

$$\text{var}(aX+b)=a^2\text{var}(X) \tag{2-31}$$

由性质(3)和性质(4)得到。例如，$\text{var}(5X+9)=25\,\text{var}(X)$。

(6) 如果 X 与 Y 相互独立，a, b为常数，那么，

$$\text{var}(aX+bY)=a^2\text{var}(X)+b^2\text{var}(Y) \tag{2-32}$$

由前面的性质可得到。例如，

$$\text{var}(3X+5Y)=9\text{var}(X)+25\,\text{var}(Y)$$

2.7.3 协方差

期望值和方差是描述单变量概率密度函数最常用的数字特征——前者给出了中心值，后者描述了单个值是围绕中心分布的程度。但是，一旦超出单变量概率密度函数的情况 (比如例2.14)，就需要考虑除了期望值和方差之外，其他多维随机变量概率分布函数的数字特征(characteristics of multivariate DDFs)，比如协方差(covariance)，相关系数(correlation)。

令随机变量 X 和 Y 的期望分别为u_x，u_y，其协方差为：

$$\begin{aligned}\text{cov}(X,Y)&=E[(X-u_x)(Y-u_y)]\\&=E(XY)-u_xu_y\end{aligned} \tag{2-33}$$

式(2-33)表明，协方差是一种特殊形式的期望值，它是两变量同时变动的度量，我们从下面的例中可以看到。

式(2-33)给出了协方差的定义式，通常用下式来计算协方差，这里假设 X与Y是离散型随机变量。

$$\begin{aligned}\text{cov}(X,Y)&=\sum_X\sum_Y(X-u_x)(Y-u_y)f(X,Y)\\&=\sum_X\sum_Y XYf(X,Y)-u_xu_y\end{aligned} \tag{2-34}$$

注意在这个式子中有两个求和符号，这是因为协方差是对两个变量的所有取值求和。对于连续型随机变量可用类似的公式计算协方差，只不过用积分符号代替求和符号。

一般而言，两随机变量的协方差可正可负。如果两个变量同方向变动 (比如，一个变量增加，另一个变量也增加)，则协方差为正，比如例 2.21；如果两个变量反方向变动 (比如，一个变量增加，另一个变量却减少)，则协方差为负。

例2.21

计算例2.14中变量 X (债券等级)和变量 Y (债券收益)的协方差。利用式(2-34)。首先用表2-3给出的数据计算 $\sum_x \sum_y XYf(X,Y)$：

$$\begin{aligned}\sum_x \sum_y XYf(X,Y) &= (8.5)(1)(0.26)+(8.5)(2)(0.10)+(8.5)(3)(0.00)\\ &\quad +(11.5)(1)(0.04)+(11.5)(2)(0.28)+(11.5)(3)(0.04)\\ &\quad +(17.5)(1)(0.0)+(17.5)(2)(0.02)+(17.5)(3)(0.26)\\ &= 26.54\end{aligned}$$

由例2.18和例2.19可知，$E(X)=u_x=2.0$，$E(Y)=u_y=12.10$。因而有：

$$\mathrm{cov}(X,Y)=26.54-(2.0)(12.10)=2.34$$

即债券等级与债券收益的协方差为正。这个结果我们不难理解，看看我们是如何编制债券等级的：等级3代表了风险最大的债券(也即B级)。因此，风险越高，期望收益就越大。

协方差的性质

下面来看协方差的一些重要性质，在以后的学习中，我们会发现在回归分析中，这些性质相当重要。

(1) 若随机变量 X，Y 相互独立，则其协方差为零。

很容易验证这条性质。如果两变量是独立的，则：

$$E(XY)=E(X)E(Y)=u_x u_y \tag{2-35}$$

把上式带到式(2-33)中，得到两个变量的协方差为零。

(2) $\mathrm{cov}(a+bX, c+dY)=bd\,\mathrm{cov}(X,Y)$ (2-36)

其中，a，b，c，d 为常数。

(3) $\mathrm{cov}(X,X)=\mathrm{var}(X)$ (2-37)

即变量与其自身的协方差就是变量的方差，可以从方差和协方差的定义加以验证。

2.7.4 相关系数

在前面讨论的债券等级一例中，我们得到债券等级与债券收益之间的协方差为 +2.34，表明这两个变量正相关。但是计算结果 2.34并未告诉我们两变量之间的正相关程度有多大。如果我们关注的是两变量的相关程度，则可用 (总体)相关系数来刻画两变量之间的相关程度。相关系数定义如下：

$$\rho=\frac{\mathrm{cov}(X,Y)}{\sigma_x\sigma_y} \tag{2-38}$$

其中希腊字母ρ代表相关系数。

从式(2-38)很容易看出，两变量的相关系数等于它们的协方差与其各自的标准差之比。相关系数是刻画两个随机变量线性相关程度的一个数字特征，即两变量之间线性相关程度有多大。

1. 相关系数的性质

相关系数有如下一些重要性质：

(1) 与协方差相同，相关系数可正可负。如果两变量之间的协方差为正，则其相关系数为正，反之，若两变量之间的协方差为负，则其相关系数为负。简言之，相关系数与协方差同号。

(2) 相关系数介于-1到1之间。用符号表示为：

$$-1 \leq \rho \leq 1 \tag{2-39}$$

如果相关系数为1，则表示两变量完全正相关，如果相关系数为－1，则表示两变量完全负相关。通常ρ介于－1，1之间。

图2-7　给出了相关系数的一些典型图案。

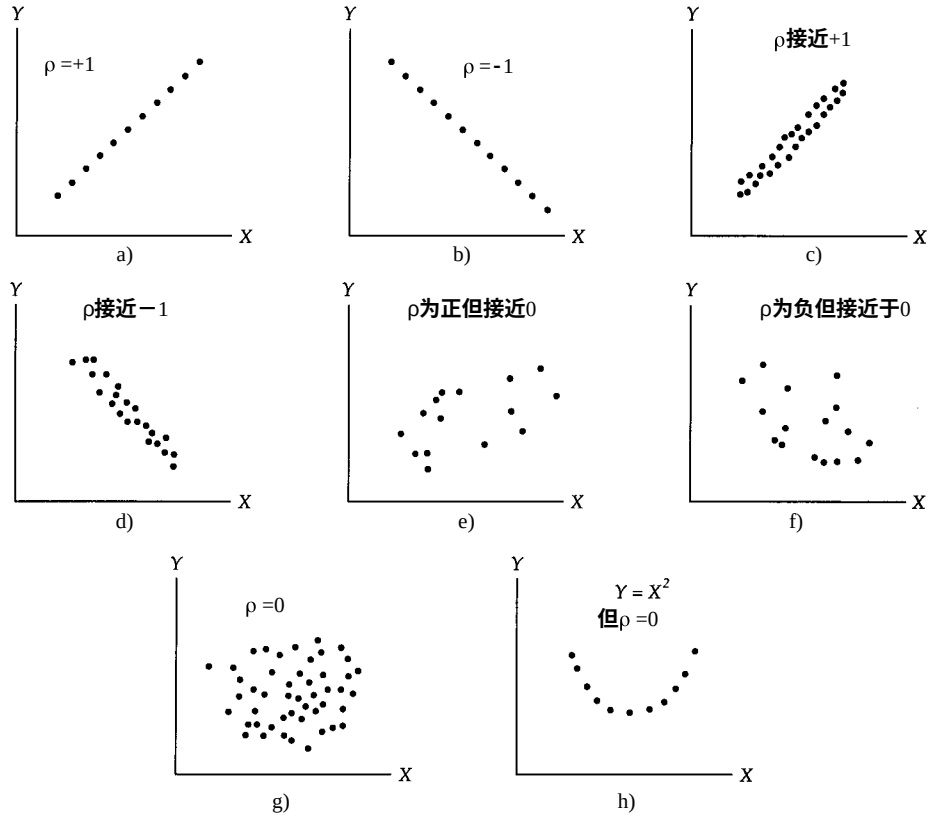


图2-7　相关系数 ρ 的典型图形

例2.22

仍以债券等级一例为例，在例 2.21中我们得到债券等级与债券收益之间的协方差为2.34。从表2-4中给出的各变量的概率密度函数，读者能够很容易求得 X (债券等级)和 Y (债券收益)的标准差分别为0.77和3.60。因此，本例中的相关系数为：

$$\rho = \frac{2.34}{(0.77)(3.60)} \approx 0.844$$

注意：仅从协方差2.34，我们无法知道两变量间的协方差是高还是低。但是，根据相关系数的大小就能够进一步确认变量之间的相关程度。由于 ρ最大为1，所以观察值0.844表明两变量高度正相关，因为0.844接近于1。

相关系数的用法将在第6章回归分析中讨论。

2. 相关变量的方差

前面我们讨论了两独立变量和或差的方差的计算式。但是，如果随机变量不是独立的 (即是相关的)，情况有会怎样呢？在此情况下，有如下计算公式：

$$\text{var}(X+Y)=\text{var}(X)+\text{var}(Y)+2\text{cov}(X,Y) \tag{2-40}$$

$$\text{var}(X-Y)=\text{var}(X)+\text{var}(Y)-2\text{cov}(X,Y) \tag{2-41}$$

当然，如果两变量之间的协方差为0，则var ($X+Y$)、var ($X-Y$)均等于var (X) + var(Y)。

留给读者练习：求债券等级一例中 $X+Y$ 的方差？

式(2-40)更一般的情况是：

$$\text{var}(aX+bY)=a^2\text{var}(X)+b^2\text{var}(Y)+2ab\,\text{cov}(X,Y) \tag{2-42}$$

同理，式(2-41)更一般的情况是：

$$\text{var}(aX-bY)=a^2\text{var}(X)+b^2\text{var}(Y)-2ab\,\text{cov}(X,Y) \tag{2-43}$$

其中，a , b 为常数。

例如，

$$\begin{aligned}\text{var}(5X\pm 7Y)&=25\text{var}(X)+49\text{var}(Y)\pm 2(35)\text{cov}(X,Y) \\ &=25\text{var}(X)+49\text{var}(Y)\pm 70\,\text{cov}(X,Y)\end{aligned} \tag{2-44}$$

2.7.5 条件期望值

在回归分析中，另一个特别重要的概念是条件期望值(conditional expectation)，它与前面讨论的随机变量的期望值 —— 称为非条件期望值(unconditional expectation) —— 不同。下面我们来解释这两个概念的差别。

回到债券等级一例(例2.14)，在该例中，债券等级 X 可取值1，2，3，债券收益 Y 可取值8.5%，11.5%，17.5%。X 的期望值是多少呢？从表2-4得其期望值为2.0，同样可得到 Y 的期望值为12.10%。这些期望值称为非条件期望值。

但现在考虑这样一个问题：已知债券等级为 1，求债券收益的均值？换一种表达方式，即为，给定 $X=1$， 求 Y 的条件期望值？更专业地，求 $E(Y\mid X=1)$？这就是 Y 的条件期望值，类似地，求 $E(X\mid Y=8.5)$？即已知债券收益为8.5%，求 X 的条件期望值？

从上面的讨论中，很清楚地看到，在计算随机变量的非条件期望值时，不必考虑其他任一变量的信息，但是在计算条件期望值时，必须考虑。

下面给出条件期望值的定义：

$$E(X\mid Y=y)=\sum_X Xf(X\mid Y=y) \tag{2-44}$$

这里，给出的是离散型随机变量的条件期望值计算公式。$f(X\mid Y=y)$是变量 X 的条件概率密度函数，$\sum_X$ 表示对所有的 X 求和。与(2-44)相比，前面讨论过的 $E(X)$称为非条件期望值。把式(2-44)与式(2-18)作比较，很清楚地看到，条件期望值 $E(X\mid Y=y)$与非条件期望值的计算公式很相像，所不同的是在条件期望值中的密度函数是条件概率密度函数，而不像在非条件期望值中用的是非条件期望值。

同样，我们给出变量Y的条件期望值：

$$E(Y\mid X=x)=\sum_Y Yf(Y\mid X=x) \tag{2-45}$$

我们用一个具体例子来说明如何计算条件期望值。

计算表明，在给定 $X=1$条件下，Y 的条件期望值为8.89%，而在前面计算过，Y 的非条件期望值为12.10%。对于这个结果我们不应感到惊讶 —— 在已知债券等级是最高一级 (三个B)的条件下，我们预期的债券收益当然会低一些。正如前面看到的条件概率密度函数通常不同于边缘概率密度函数一样，一般条件期望值也不同于非条件期望值。

在结束有关概率分布特性的讨论之前，我们再来看另外两个概念 —— 概率密度函数的偏度(skewness)和峰度(kurtosis)。偏度和峰度是用于描述概率密度函数形状的数字特征。偏度 (S)是

对称性的度量，峰度(K)是一概率密度函数高低或胖瘦的度量，见图 2-8。

例2.23

在债券等级一例中，求 $E(Y \mid X=1)$？即求在已知债券等级为1的条件下，债券收益的期望值。

利用式(2-45)，得到[1]，

$$E(Y \mid X=1)=\sum Y f(Y \mid X=1)$$
$$=8.5 f(Y=8.5 \mid X=1)+11.5 f(Y=11.5 \mid X=1)+17.5 f(Y=17.5 \mid X=1)$$
$$=8.5(0.87)+11.5(0.13)+17.5(0)$$
$$=8.89$$

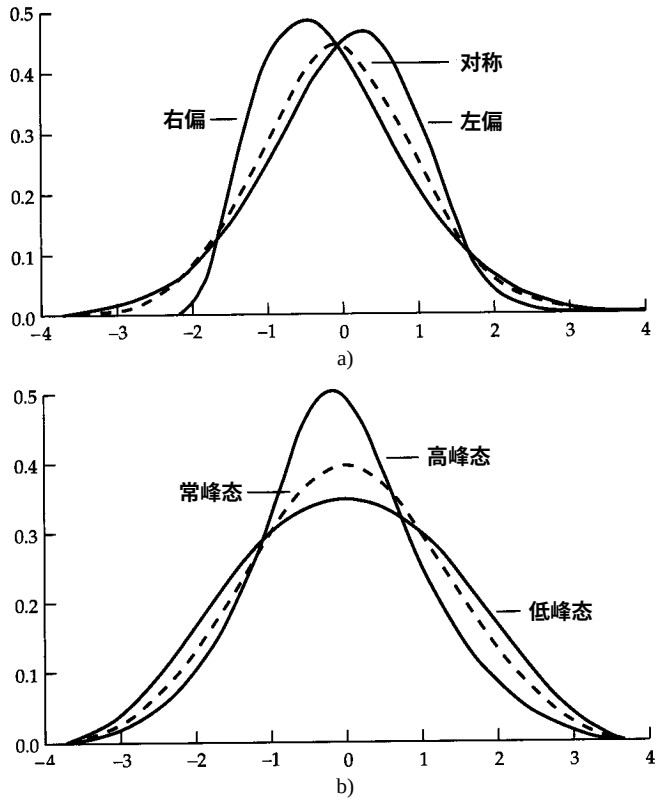


图　2-8
a)偏度　b)峰度

在度量偏度和峰度之前，首先需要了解概率密度函数的三阶矩与四阶矩。我们已经知道随机变量 X 的一阶矩是 $E(X)=u_x$，即 X 的均值，二阶中心矩是 $E(X-u_x)^2$(即 X 的方差)。按照这种方式，三阶中心矩和四阶中心矩表示为：

三阶中心矩：$E(X-u_x)^3$

四阶中心矩：$E(X-u_x)^4$

一般 r 阶中心矩可表示为：

1　$f(Y=8.5|X=1)=f(Y=8.5,X=1)/f(X=1)=0.87$。即条件概率密度函数等于联合概率密度函数与边缘概率密度函数之比，参见等式(2-16)。

r 阶中心矩： $E(X-u_x)^r$

给出这些定义以后，通常运用下列公式计算偏度和峰度 ：

$$S=\frac{[E(X-u_x)^3]^2}{[E(X-u_x)^2]^3} \tag{2-46}$$

$$=\frac{\text{三阶矩的平方}}{\text{两阶矩的立方}}$$

因为，对于对称的概率密度函数，其三阶矩为零，因此这样的一个概率密度函数，其偏度 S 为零。一个最重要的例子就是正态分布，我们将在下一章详细讨论。如果偏度 S 的值为正，则其概率密度为正偏或右偏；如果 S 的值为负，则其概率密度为负偏或左偏。

$$K=\frac{E(X-u_x)^4}{[E(X-u_x)^2]^2} \tag{2-47}$$

$$=\frac{\text{四阶矩}}{\text{两阶矩的平方}}$$

概率密度函数的峰度 K 小于3时，成为低峰态的(胖的或短尾的)，峰度 K 大于3时，称为尖峰态的(瘦的或长尾的)，见图2-8。正态分布的峰度 K 为3，这样的概率密度函数称为常峰态的。

在随后的章节中将大量使用正态分布，因此，对其偏度 S (=0)和峰度 K (=3)的了解将有助于比较其他的概率分布函数。

直接扩展式(2-27)，即可得到三阶矩和四阶矩的计算公式，

$$\text{三阶矩：}\ \sum(X-u_x)^3 f(X) \tag{2-48}$$

$$\text{四阶矩：}\ \sum(X-u_x)^4 f(X) \tag{2-49}$$

其中 X 是离散型随机变量。对于连续型随机变量，我们将求和符号换成积分符号(∫)即可。

例2.24

考虑表2-6给出的概率密度函数。已知期望值和方差分别为 3.5，2.916 7。三阶中心矩和四阶中心矩的计算如下：

X	$f(X)$	$(X-u_x)^3 f(X)$	$(X-u_x)^4 f(X)$
1	1/6	$(1-3.5)^3(1/6)$	$(1-3.5)^4(1/6)$
2	1/6	$(2-3.5)^3(1/6)$	$(2-3.5)^4(1/6)$
3	1/6	$(3-3.5)^3(1/6)$	$(3-3.5)^4(1/6)$
4	1/6	$(4-3.5)^3(1/6)$	$(4-3.5)^4(1/6)$
5	1/6	$(5-3.5)^3(1/6)$	$(5-3.5)^4(1/6)$
6	1/6	$(6-3.5)^3(1/6)$	$(6-3.5)^4(1/6)$
	求和：	0	14.732

根据偏度和峰度的定义，读者可以验证这里的偏度为 0(很令人惊讶吗？)，峰度为1.731 7。因此，虽然这个概率密度函数以均值呈中心对称，但它却是低峰态分布，即比正态分布略“胖”一些，我们可从图 2-6明显地看出来。

2.8 从总体到样本

为了计算概率分布的特性，比如期望值、方差、协方差、相关系数、条件期望值等等，我们显然需要概率密度函数，也即整个样本空间或总体 (即概率密度函数)。因此，要想知道在某一时刻，居住在纽约市所有居民的平均收入，显然需要知道纽约全部人口。虽然在理论上，某一时点上的居住在纽约的居民是有限的，但在实际中很难收集总体中每一个成员的信息 (用概

率的语言，也即结果)。实践中我们所能做到的是从总体中抽取一个“有代表性的”或“随机的”样本，然后计算抽样样本的人均收入。[1]

但是，从样本中得到的平均收入等于总体真实的平均收入吗？很可能不同。类似的，如果我们从抽样总体中计算收入的方差，那么，它等于总体真实的方差吗？同样，很可能不同。

那么，如果仅仅有来自于总体的一两个样本，怎样才能知道总体的数字特征呢？比如期望，方差等等。通览本书，你会发现，实际上，我们都无一例外地依赖于来自某个总体的一个或多个样本。

这个重要问题的答案将是本书第4章讨论的重点。但首先我们必须求出与各种总体数字特征相对应的样本矩(sample moments)。

2.8.1 样本均值

随机变量 X 代表某汽车销售商每天销售汽车的数量。若 r.v X 服从某一概率密度函数。而且，我们想求每月前十天，该汽车销售商出售汽车的平均数量 (即期望值)。假设该汽车销售商已从业十年，但在过去的十年里，没有时间细看每月前十天销售数量。若该销售商从过去的数据中随机抽取某月销售量，并记下该月前十天的汽车销售数量：9，11，11，14，13，9，8，9，14，12。这就是一个包括十个样本值的样本。注意该汽车销售商共有 120个月的数据，如果他决定抽取另外一个月，则可能得到另外十个不同值。

如果该销售商把十个样本值相加求和再除以10(样本容量)，即为样本均值。

随机变量的样本均值通常用符号 $\bar{X}$ 表示，定义如下：

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} \tag{2-50}$$

其中 $\sum_{i=1}^{n} X_i$ 表示从1到 n 对所有的 X 值相加，n 为样本容量。

上面定义的样本均值就是总体均值 (期望 $E(X)$)的估计量(estimator)。估计量可以简单地理解为估计总体(比如总体均值)的规则或公式。在第3章中，我们将讨论怎样用样本均值 $\bar{X}$ 估计期望值 $E(X)$。

在本例中，样本均值为：

$$\bar{X} = \frac{9+11+11+\cdots+12}{10} = 11$$

我们称样本均值是总体均值的估计值(estimate)。估计值简单地说就是估计量的取值，例如在本例中是11。在这个例子中，每月前十天汽车销售的平均数量为 11，但这个值并不一定等于 $E(X)$。要计算 $E(X)$，需要考虑其他119个月前十天的汽车销售量。简言之，我们需要考虑整个的概率密度函数。但是，在第3章中你会看到，一般地，样本估计值 (比如11)很好的近似了真实的 $E(X)$。

2.8.2 样本方差

在上例中给出的10个样本值并不全都等于样本均值 11。这种变异性可用样本方差 (S_x^2)来度量。它是总体方差 σ_x^2 的估计量。样本方差(sample variance)的定义如下：

$$S_x^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1} \tag{2-51}$$

1 随机样本的准确定义将在后面给出。(参见第3章)

即样本方差等于每个 X 与其均值差的平方除以 $n-1$再求和。[1] $(n-1)$称为自由度，它的准确定义将在第3章给出。[2] S_x^2 的正的平方根 S_x 称为样本标准差(sample standard deviation,样本s.d)。

由上例中给出的10个样本值求得其样本方差为：

$$S_x^2 = \frac{(9-11)^2+(11-11)^2+\cdots+(12-11)^2}{9}$$

$$= 4\ 4\ /\ 9 \approx 4.89$$

样本标准差 $S_x = \sqrt{4.89} = 2.21$ 。注意4.89是总体方差的估计值， 2.21是总体标准差的估计值。再一次强调，估计值是样本估计量的取值。

2.8.3 样本协方差

例2.25

设有两变量X(股票价格)和Y(消费者价格)构成的二元总体。进一步假设从该二元总体中得一随机样本,见表2-8的第一、第二列。在这个例子中,股票价格用道-琼斯指数均值来度量,消费者价格用消费者价格指数(CPI)度量。表中其他各值我们随后讨论。

类似式(2-33)总体协方差的定义,两随机变量之间的样本协方差(sample covariance)的定义如下：

$$\text{样本 cov}(X,Y) = \frac{\sum (X_i-\bar{X})(Y_i-\bar{Y})}{n-1} \tag{2-52}$$

即样本协方差为两随机变量与其各自的 (样本)均值求差再除以自由度,$(n-1)$，(如果样本容量足够大，可用 n 作除数)然后对其差积求和。式(2-52)定义的样本协方差是总体协方差的估计量。本例中给出的样本协方差的数值即为总体协方差的估计值：

样本协方差 $\text{cov}(X,Y)=63\ 294/9=7\ 026.20$

因此，本例中的股票价格与消费者价格的协方差为正。有些分析家认为投资股票可以预防通货膨胀，也就是说，当通货膨胀加剧时，股票的价格也会上升。虽然对于这一结论缺乏经验的证据，但1980～1989年期的确是这样的。

表2-8 1980-1989年道-琼斯指数均值(X)与消费者价格指数(Y)的样本协方差及样本相关系数

Y	X	$(Y-\bar{Y})(X-\bar{X})$	
(1)	(2)	(3)	
891.4	82.4	(891.4－1 504.4)	(82.4－104.64)
932.42	90.9	(932.42－1 504.4)	(90.9－1 504.4)
884.36	96.5	—	—
1 190.34	99.6	—	—
1 178.48	103.9	—	—
1 328.23	107.6	—	—
1 792.76	109.6	—	—
2 275.99	113.6	—	—
2 060.82	118.3	—	—
2 508.91	124.0	—	—
15 044	1 046.4	≈63 234	
$\bar{Y}$ =15 044/10=1 504.4	样本方差 var(Y)=368.870		
$\bar{X}$ =1 046.4/10=104.64	样本方差 var(X)=161.18		

资料来源：Data on Y and X are from the *Economic Report of the President,* 1996, Tables B-91, p.384, and B-56, p.343, respectively.

1 如果样本容量很大，式(2-51)中的分子可以用n除，而不是$n-1$。

2 在式(2-51)中，用除数$n-1$的原因是，式(2-51)给出了真实的σ_x^2的一个无偏估计量。也就是说，如果重复使用式(2-51)，平均而言，根据式(2-51)计算的样本方差将等于真实的总体方差。估计量是无偏的定义我们将在第4章给出。

2.8.4 样本相关系数

式(2-38)定义了两随机变量的总体相关系数。类似地，样本相关系数定义如下，通常用符号 r 表示样本相关系数。

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})/(n-1)}{S_X S_Y}$$
$$=\frac{\text{样本}\operatorname{cov}(X,Y)}{\text{s.d}(X)\text{s.d}(Y)} \tag{2-53}$$

因此，样本相关系数(sample correlation)与总体相关系数有相同的性质；它们均位于- 1与1之间。

由表2-8给出的数据，读者很容易计算出 X ，Y 的样本标准差，然后求出ρ的估计值 r ：

$$r=\frac{7\,026.20}{(12.696)(607.40)}$$
$$=0.9111$$

所以，在这个例子中，股票价格与消费者加价格高度相关，因为计算出的相关系数接近于1。

2.8.5 样本偏度与样本峰度

根据式(2-46)和式(2-47)，用样本三阶矩和四阶矩来计算样本偏度与峰度。样本三阶矩 (与样本方差的计算公式相对照)为：

$$\frac{\sum(X-\overline{X})^3}{n-1} \tag{2-54}$$[1]

样本四阶矩为：

$$\frac{\sum(X-\overline{X})^4}{n-1} \tag{2-55}$$

利用表2-8给出的数据，读者可计算出样本三阶矩和四阶矩，然后可以验证道 -琼斯指数均值的样本偏度和峰度分别为 0.467 3和1.545 6，表明道-琼斯指数均值的分布是正偏的，比正态分布略“胖”一些。

2.9 小结

本章介绍了概率、随机变量、概率密度、概率密度的数字特征 (矩)等一些基本概念。我们的目的并不是教授统计学，而是复习和回顾在本书后面讨论中所必须了解的一些基本统计概念，因此，对这些概念的讨论更多的是直观的。

本章介绍了一些重要公式。这些公式告诉我们如何计算随机变量的概率以及如何估计概率分布的数字特征，比如期望值 (均值)、方差、协方差、相关系数、条件期望值等等。在介绍公式的同时，我们仔细地区分了总体矩与样本矩，并给出相应的计算公式。因此，随机变量的期望值 $E(X)$ 是总体矩，即是知道总体的所有取值情况下 X 的均值。另一方面，$\overline{X}$ 是样本矩，即它是来自于总体的一个样本，并非全部总体的 X 的均值。在统计学中，总体和样本的两分性非常重要，因为在许多实际运用过程中，我们仅有来自于总体的一两个样本，而且常常是通过这些样本矩来推断总体矩。如何利用样本对总体进行推断，我们将在随后的两章中讨论。

---

1 如果样本足够大，可用除数 n 代替 $n-1$

参考文献

正如本章前言部分提到的那样，本章只是对统计学的一些基本概念做一个简要和直观的介绍，它并不是统计学基本教程。因此，建议读者手边能有一两本好的统计学教科书。下面介绍一些参考文献：

1. Newbold,Paul:*Statistics for Business and Economics*,Prentice-Hall,Englewood Cliffs,N.J.

这本书提供大量的实例，没有复杂的数学计算。

2. Hoel,Paul G.: *Introduction to Mathematial Statistcs*,Wiley,New York.

这本书简单介绍了数理统计学的各方面的知识。

3. Mood,Alexander M., Graybill,Franklin A., and Boes,Duane C.:*Introduction to the Theory of Statistics*,McGraw-Hill,New York,1974.

这是一本标准的高级统计学教材。

4. Mosteller,F.,Rourke,R.,and Thomas ,G.:*Probability with Statistical Applications*,Addision-Wesley,Reading,Mass.

5. DeGrott,Morris H.:*Pobability and Statistics*(2nd ed.),Addison-Wesley,Reading,Mass.

习题

2.1 解释概念

(a) 样本空间 (b) 样本点 (c) 事件 (d) 互斥事件 (e) 概率密度函数 (f) 联合概率密度函数 (g) 边缘概率密度函数 (h) 条件概率密度函数 (i) 统计独立性

2.2 事件 A，B 能否既为互斥事件同时又相互独立？

2.3 什么是概率密度函数的矩？最常用的矩有哪些？

2.4 解释概念

(a) 期望值 (b) 方差 (c) 标准差 (d) 协方差 (e) 相关 (f) 条件期望值

2.5 解释概念

(a) 样本均值 (b) 样本方差 (c) 样本标准差 (d) 样本差协方差 (e) 样本相关系数

2.6 为什么说区分总体矩和样本矩很重要？

2.7 按照(a)填空。

(a) 期望值或均值是集中趋势的度量。 (b) 方差是.....的度量。 (c) 协方差是.....的度量。 (d) 相关系数是.....的度量。

2.8 随机变量 X 的均值为$50,标准差为$5,那么，方差为$25。对吗？为什么？

2.9 判断正误并解释原因。

(a) 虽然随机变量的期望值可正可负，但其方差总为正。 (b) 两变量的协方差与相关系数同号。 (c) 一个随机变量的条件期望值和非条件期望值意义相同。 (d) 若两变量相互独立，则其相关系数必定为零。 (e) 若两变量的相关系数为零，则它们相互独立。

2.10 下列各式代表什么？

(a) $\sum_{i=1}^{4} x^{i-1}$ (b) $\sum_{i=1}^{6} ay_i$ a为常数 (c) $\sum_{i=1}^{2} (2x_i + 3y_i)$ (d) $\sum_{i=1}^{3}\sum_{j=1}^{2} x_i y_i$

(e) $\sum_{i=1}^{4} (i + 4)$ (f) $\sum_{i=1}^{3} 3^i$ (g) $\sum_{i=1}^{10} 2$ (h) $\sum_{i=1}^{3} (4x^2 - 3)$

2.11 用求和符号表示下列各式：

(a) $x_1+x_2+\cdots+x_5$ (b) $x_1+2x_2+3x_3+4x_4+5x_5$ (c) $(x_1^2+y_1^2)+(x_2^2+y_2^2)+\cdots+(x_k^2+y_k^2)$

2.12　利用等式 $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ 求下列各式。

(a) $\sum_{k=1}^{500} k$　(b) $\sum_{k=10}^{100} k$　(c) $\sum_{k=10}^{100} 3k$

2.13　可以证明等式 $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$ 成立，利用该等式计算下列各式。

(a) $\sum_{k=1}^{10} k^2$　(b) $\sum_{10}^{20} k^2$　(c) $\sum_{k=1}^{10} 4k^2$

2.14　随机变量的概率密度函数如下：

X	$f(X)$	X	$f(X)$
0	b	3	$4b$
1	$2b$	4	$5b$
2	$3b$		

(a) 求 b？为什么？　(b) 求 $P(X \leq 2)$；$P(X \leq 3)$；$P(2 \leq X \leq 3)$？　(c) 求 X 的期望 $E(X)$？　(d) 求 X 的方差 $\mathrm{var}(X)$？

2.15　随机变量的变差系数用 V 表示，其定义为：

$$V = \frac{E(X)}{\sigma_x} = \frac{X\text{的期望}}{X\text{的标准差}}$$

计算习题2.14中的 V 值。如果 X 的单位为美元，那么 V 的单位是什么？(注：实际中，V 通常以百分比的形式表示)。

2.16　下表给出了某项投资1年后预期的回报率及相应概率。

(a) 求投资回报率的期望值？

(b) 求投资回报率的方差和标准差？

(c) 求投资回报率的偏度系数和峰度系数？

(d) 求累积概率密度函数及回报率小于等于10%的概率？

报酬率 X (%)	$f(X)$
−20	0.10
−10	0.15
10	0.45
25	0.25
30	0.05
总计	1.00

2.17　下表给出了变量 X，Y 的联合概率密度函数

Y \ X	1	2	3
1	0.03	0.06	0.06
2	0.02	0.04	0.04
3	0.09	0.04	0.04
4	**0.06**	**0.12**	**0.12**

(a) 求 X 和 Y 的边缘(非条件)分布 $f(X)$，$f(Y)$？

(b) 求条件概率密度函数，$f(X \mid Y)$，$f(Y \mid X)$？

(c) 求条件期望值 $E(X \mid Y)$，$E(Y \mid X)$？

(d) 求 $E(X)$，$E(Y)$？

(e) 求出的条件期望值与非条件期望值相同吗？

(f) X 与 Y 相互独立吗？你是如何判断的？

2.18　下表给出了随机变量 X，Y 的联合概率密度函数，其中 X 表示投资项目A的1年期报

酬率，Y 表示项目B的1年期报酬率。

Y (%) \ X (%)	−10	0	20	30
20	0.27	0.08	0.16	0.00
50	0.00	0.04	0.10	0.35

(a) 计算项目A的期望值的报酬率。

(b) 计算项目B的期望值的报酬率。

(c) 变量 X 与 Y 相互独立吗？(提示：$E(XY)=\sum_{X=1}^{4}\sum_{Y=1}^{2}X_iY_if(X_iY_i)$)

2.19 已知 $E(X)=8$，$\text{var}(X)=4$，求下列各式的期望值及方差？

(a) $Y=3X+2$ (b) $Y=0.6X-4$ (c) $Y=X/4$ (d) $Y=aX+b$, 其中 a , b 为常数

2.20 证明：$\text{var}(X)=E[X-E(X)]^2=E(X)^2-[E(X)]^2=E(X)^2-u_X^2$

(a) $\text{cov}(X,Y)=E[(X-u_X)(Y-u_Y)]=E(XY)-u_Xu_Y$

其中，$u_X=E(X)$，$u_Y=E(Y)$

如何用文字表述上列各式？

2.21 利用等式(2-38)给出的总体相关系数的定义，证明：

$$\text{var}(X+Y)=\text{var}(X)+\text{var}(Y)+2\rho\sigma_X\sigma_Y \tag{2-40a}$$

$$\text{var}(X-Y)=\text{var}(X)+\text{var}(Y)-2\rho\sigma_X\sigma_Y \tag{2-41a}$$

2.22 考虑上题给出的等式(2-40a)。令 X 代表债券的收益率(比如说，IBM)，Y 代表另一种债券的收益率(比如说，一般物品)。令 $\sigma_X^2=16$，$\sigma_Y^2=9$，$\rho=-0.8$。求 $X+Y$ 的方差。它比 $\text{var}(X)+\text{var}(Y)$ 大还是小。在此例中，是否同等投资于两种债券(即多样化)比完全投资于其中一种的收益大？这个问题是现代有价证券理论的精华所在。(参见Richard Brealey, Steward Myers, Priciples of Corporate finace,McGraw-Hill，New York,1981)

2.23 100个人中，有50人是民主党成员，40人是共和党成员，10人是无党派人士。在这三类人中阅读《Wall Street》杂志的比例分别是30%，60%，40%。若有一人正在读该杂志，问这个人是共和党成员的概率是多少？

2.24 表2-9给出了美国1979至1988年新成立公司与破产公司的数据。

表2-9 美国1979-1988年期间新成立公司(X)与破产公司(Y)数目

年份	Y	X	年份	Y	X
1979	524 565	7 564	1984	634 911	52 078
1980	533 520	11 742	1985	662 047	57 253
1981	581 242	16 794	1986	702 738	61 616
1982	566 942	24 908	1987	685 572	61 622
1983	600 400	31 334	1988	685 095	57 099

资料来源：*Economic Report of the President*,1990,Table C-94,p.402.

(a) 求新成立公司数目的均值和方差？ (b) 求破产公司数目的均值是和方差？ (c) 求 X 与 Y 的协方差与相关系数？ (d) 这两个变量相互独立吗？ (e) 如果两个变量相关，是否可以认为一个变量是另一个变量的“原因”。即是新公司的进入导致原有公司的破产，还是原有公司的破产导致新公司的进入？

2.25 在例2.14中，求 $\text{var}(X+Y)$。你如何解释这个方差？

2.26 参见习题1.6中的表1-2，

(a) 计算S&P500与CPI的协方差以及CPI与3月期国债利率的协方差，这是样本协方差还是总体协方差？ (b) 计算S&P500与CPI的相关系数以及CPI与3月期国债利率的相关系数，先验地，你认为这些相关系数是正还是负，为什么？ (c) 如果CPI与3月期国债利率正相关，是否意味着以CPI来衡量的通货膨胀是较高国债利率的原因？

2.27 参见习题1.6中的表1-3，令ER代表德国马克对美元的汇率(即一美元兑换多少德国马克)，RPR代表美国消费者价格指数与德国消费者价格指数之比。你预期 ER与RPR是正相关还是负相关？为什么？写出计算步骤。如果知道 ER与1/RPR的相关关系，你会改变答案吗？为什么？

第 3 章

一些重要的概率分布

在前面的章节中我们讲到随机变量可以用其概率密度函数的一些数字特征(或矩)来描述，比如期望值和方差。但是，由于随机变量种类繁多，因此假设知道其概率密度函数实际上是较高的要求。但在实际中，一些随机变量经常发生，因此统计学家能够确定其概率密度函数并归纳出其性质。这里，我们主要关注的是一些基本的概率密度函数。但是，在任何一本标准的统计学教科书上，你都会发现统计学家还对其他的一些概率密度函数作了仔细的研究。

本章主要讨论的4种概率分布是：

(1) 正态分布；(2) χ^2分布；(3) t分布；(4) F分布。

我们将考察上述各概率密度的主要特征、性质及其用途。读者必须掌握本章的全部内容，因为，这些概率分布是经济计量理论和实践的核心内容。

3.1 正态分布

对于连续型随机变量而言，正态分布(normal distribution)是最重要的一种概率分布，稍具统计知识的读者都会熟悉其“钟型”形状(见图 2-2)。经验表明：对于其值依赖于众多微小因素且每一因素均产生微小的或正或负影响的连续型随机变量来说，正态分布是一个相当好的描述模型。比如考虑体重这一随机变量，它就近似服从正态分布，因为遗传、骨骼结构、饮食、锻炼、新陈代谢等都对人的体重有影响，但又没有一种因素起到压到一切的主导作用。与此相类似，人的身高、考试分数等都近似地服从正态分布。

为了简便，通常用：

$$X \sim N(u, \sigma^2) \qquad (3\text{-}1)^1$$

表示随机变量 X 服从正态分布。符号～表示随机变量服从什么样的分布，N表示正态分布，括号内的参数u,σ^2称为正态分布的(总体)均值(或期望)和方差。需要指出的是：X是一个连续型随机变量，可取区间$(-\infty,+\infty)$内的任意一值。

---

1 正态变量的概率密度函数：

$$f(X)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{X-u}{\sigma}\right)^2\right\}$$

其中,exp{}表示以e为底的指数形式，e=2.718 28，π=3.141 59。μ和σ^2分别是正态分布的参数,均值和方差。

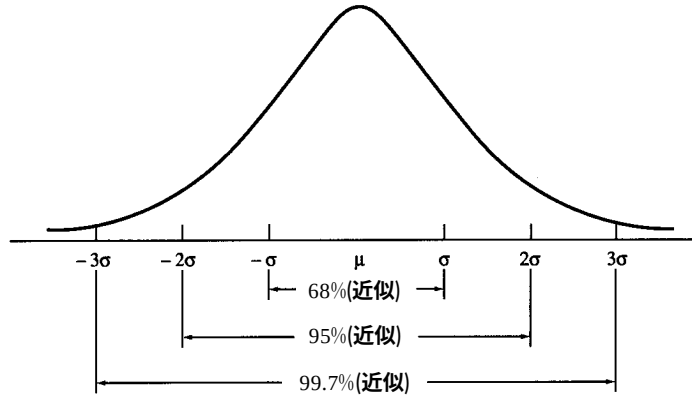


图3-1　正态曲线下的区域

3.1.1　正态分布的性质

(1) 正态分布曲线(见图2-2)以均值u为中心，对称分布。

(2) 正态分布的概率密度函数呈中间高、两边低，在均值 u处达到最高，向两边逐渐降低，即随机变量在远离均值处取值的概率逐渐变小。例如，某人身高超过 7.5英寸的概率非常小。

(3) 正态曲线下的面积约有 68%位于$u \pm \sigma$两值之间；约有95%的面积位于$u \pm 2\sigma$之间；而约有99.7%的面积位于$u \pm 3\sigma$之间。这些区域可用作概率的度量。

(4) 正态分布可由两个参数u，σ^2来描述，即一旦知道u，σ^2的值就利用脚注1给出的数学公式求得随机变量 X 落于某一区间的概率。幸运的是，我们不必利用这个公式进行如此繁琐的计算，可以根据附录A中的表A-1很容易查到这些概率值。随后解释如何使用该表。

(5) 两个(或多个)正态分布随机变量的线性组合仍服从正态分布 —— 在经济计量学中，这是正态分布的一个特别重要的性质。

令：

$$X \sim N(u_X, \sigma_X^2)$$
$$Y \sim N(u_Y, \sigma_Y^2)$$

假定X和Y相互独立。[1]

现在考虑这两个变量的线性组合：$W = aX + bY$

其中a、b为常数(例如：$W = 2X + 4Y$)。则，

$$W \sim N(u_W, \sigma_W^2) \tag{3-2}$$

其中，

$$u_W = (au_X + bu_Y) \tag{3-3}$$
$$\sigma_W^2 = (a^2\sigma_X^2 + b^2\sigma_Y^2) \tag{3-4}$$

注意：在式(3-3)中，利用了第2章讨论过的独立随机变量期望和方差的一些性质(见2.7节)。[2]

1　如果两个变量的联合概率密度函数等于其各自边缘概率密度函数之积，那么，这两个变量相互独立。即，对于所有的X和Y，$f(XY)=f(X)f(Y)$成立。

2　需要指出的是：如果X、Y服从正态分布，但不相互独立，则它们的线性组合W仍服从正态分布，期望与式(3-3)给出的期望相同，但方差为：

$$\sigma_W^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\mathrm{cov}(X,Y)$$

与此同时，可把式(3-2)推广到两个以上随机变量的线性组合的情形。

例3.1

令X表示在曼哈顿非商业区一花商每日出售玫瑰花数量，Y表示在曼哈顿商业区一花商每日出售玫瑰花的数量，假定X和Y服从正态分布，且相互独立，并有：$X\sim N(100, 64)$，$Y\sim N(150, 81)$，求两天内两花商出售玫瑰花数量的期望及方差？

$$W=2X+2Y$$

因此，根据式(3-3)，得

$$E(w)=E(2X+2Y)=500,$$

$$\mathrm{Var}(w)=4\mathrm{var}(X)+4\mathrm{var}(Y)=580$$

因此，W服从均值为500，方差为580的正态分布，即$W\sim N(500, 580)$。

(6) 正态分布的偏度(S)为0，峰度(K)为3。

3.1.2 标准正态分布

虽然正态分布完全可由其两个数字特征——(总体)均值和方差来描述，但是，两个正态分布可能因为期望或方差的不同，或是期望和方差均不同而相区别。(见图 3-2)

如何比较图3-2中所描绘的各种不同的正态分布呢？既然这些正态分布由于期望或方差的不同或是期望和方差均不相同而有差别，我们不妨定义一个新的变量 Z：

$$Z=\frac{X-u}{\sigma}$$

如果变量X的均值为u，方差为σ，则根据式(3-4)的定义，变量Z的均值为0，方差为1。在统计学中，我们称之为单位或标准正态变量(unit,or standard normal variable)。即若$X\sim N(u, \sigma^2)$，那么变量Z就是单位正态变量或标准正态变量，也就是说，标准正态变量的均值为 0，方差为1，用符号表示为：

$$Z\sim N(0, 1) \qquad (3\text{-}5)^1$$

因此，任一给定均值和方差的正态变量都可转化为标准正态变量，我们会发现，这种标准化会大大简化计算。

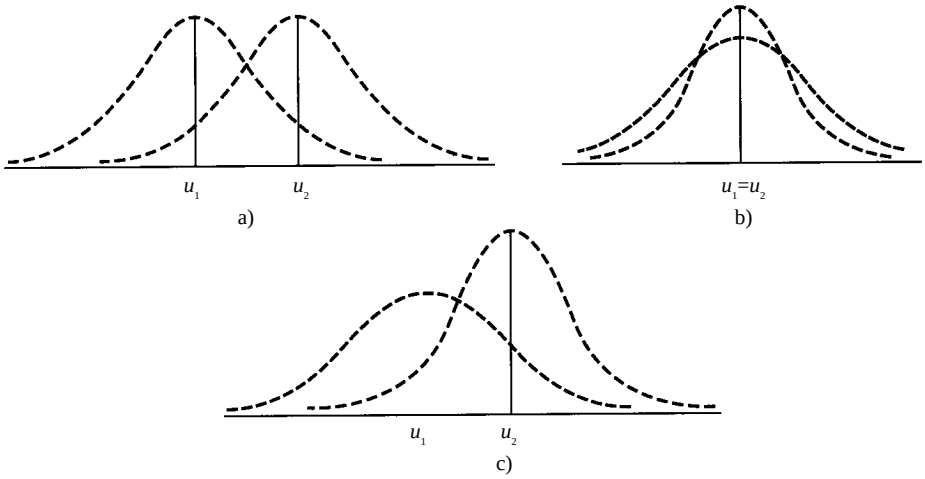


图 3-2

a) 不同均值，同方差 b) 同均值，不同方差 c) 不同均值，不同方差

1 利用正态分布的性质，很容易证明：正态变量的线性函数仍服从正态分布。

图3-3a和3-3b分别给出标准正态分布的概率密度函数和累积分布函数，（见2.5节中PDF和CDF的定义）如同任何一种累积分布函数，标准正态分布的累积分布函数给出了标准正态变量取小于或等于某一z值的概率（即，$P(Z\leq z)$），其中z为某一给定的Z值。

我们通过下面的几个具体例子来说明如何使用标准正态分布表。

例3.2

变量x表示面包房每日出售的面包量，假定它服从均值为 70、方差为9的正态分布，即$X\sim N(70, 9)$，求任给一天，出售面包数量大于 75条的概率。

$$Z=\frac{75-70}{3}\approx 1.67$$

服从标准正态分布，现求$P(Z>1.67)$。[1]

附录A中表A-1给出Z值从0到3.09的标准正态分布的累积概率密度函数。例如，从该表可知，Z位于区间(0,1.3)的概率为0.403 2(40.32%)，位于(0，2.5)的概率为0.493 8(49.38%)。由正态分布的对称性可知，Z位于区间(-1.3,0)的概率也为0.403 2，位于(-2.5,0)的概率为0.493 8。正是由于这种对称性，在标准正态分布表中一般仅给出取正值的情形。也就是说，标准正态密度函数，在z=0的左右面积均为0.5，整个面积(或概率)为1。

根据正态分布表得：

$$P(0\leq Z\leq 1.67)=0.452\ 5,$$

因此，

$$P(Z>1.67)=0.500\ 0-0.425\ 7=0.047\ 5$$

即每天出售面包的数量超过75条的概率为0.047 5。(参见图3-3a)

例3.3

继续例3.2,现假定要求每天出售面包数量小于或等于 75条的概率。

读者很容易验证这个概率为：0.500 0+0.452 5=0.952 5 (见图3-3b)。

例3.4

现求每天出售面包数量在65条与75条之间的概率。

为求此概率，先要计算

$$Z_1=\frac{65-30}{3}\approx -1.67$$

$$Z_2=\frac{75-70}{3}\approx 1.67$$

查表A-1，得，

$$P(-1.67\leq Z\leq 0)=0.452\ 5$$

$$P(0\leq Z\leq 1.67)=0.452\ 5$$

由正态分布的对称性得到，

$$P(-1.67\leq Z\leq 1.67)=0.905\ 0$$

即每天出售面包的数量介于65条与75条之间的概率约为90.5%(见图3-3a)。

1 写$P(Z\geq 1.67)$或$P(Z>1.67)$没有什么实质差别，因为，在第 2章我们已经讲过，连续型随机变量取某一特殊值(比如1.67)的概率为零。

例3.5

现求每天出售面包数量大于75条或小于65条的概率。

如果读者已经掌握前面各例，就很容易求得其概率为0.095(见图3-3a)。

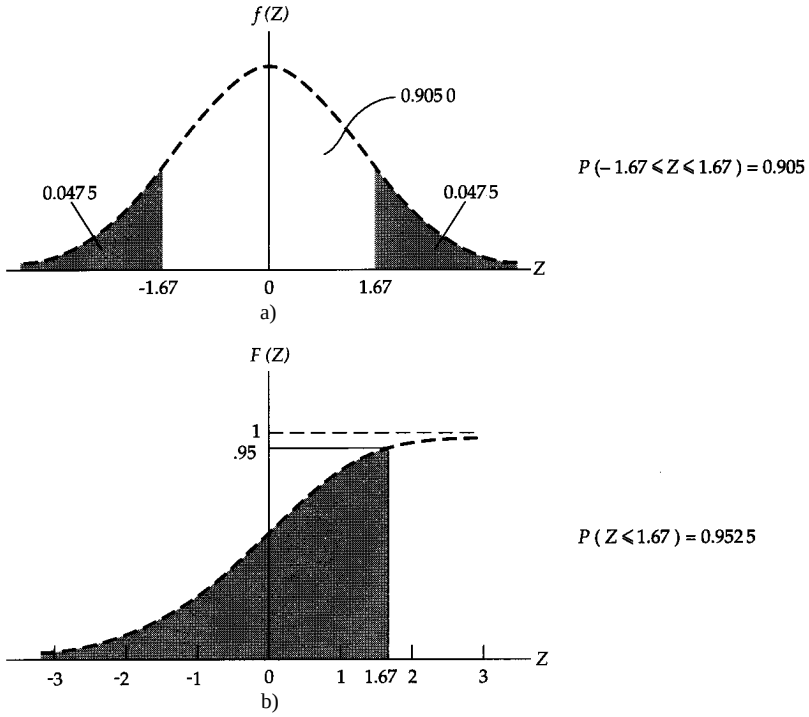


图3-3

a) 标准正态变量的PDF b) 标准正态变量的CDF

上面的例子表明：一旦知道某一正态变量的期望与方差，先将其转化为标准正态变量，然后根据正态分布表求得相应的概率。的确很神奇，仅仅通过一个标准正态分布表就能处理任一正态分布变量，而无论其均值和方差是多少。

正如前面讲到的那样，正态分布是最重要的理论概论分布，因为许多(连续型)随机变量都服从正态分布或近似正态分布，我们将在3.2节中详细讨论。在此之前，先来看在使用正态分布时遇到的一些实际问题。

表3-1 来自$N(0,1)$，$N(2,4)$的25个随机数

$N(0,1)$	$N(2,4)$	$N(0,1)$	$N(2,4)$
-0.48524	4.25181	0.22968	0.21487
0.46262	0.01395	-0.00719	-0.47726
2.23092	0.09037	-0.71217	1.32007
-0.23644	1.96909	-0.53126	-1.25406
1.10679	1.62206	-1.02664	3.09222
-0.82070	1.17653	-1.29535	1.05375
0.86553	2.78722	-0.61502	0.58124
-0.40199	2.41138	-1.80753	1.55853
1.13667	2.58235	0.20687	1.71083
-2.05585	0.40786	-0.19653	0.90193
2.98962	0.24596	2.49463	-0.14726
0.61674	-3.45379	0.94602	-3.69238
-0.32833	3.29003		

3.1.3 从正态总体中随机抽样

在统计学中，正态分布运用得十分广泛，因此，知道怎样从一正态总体中抽取一随机样本就很重要。假设我们想从一均值为0，方差为1，[即标准正态分布，$N(0，1)$]正态分布中抽取一个有25个观察值的随机样本，如何得到这样一个随机样本呢？

许多统计软件包都有从常用的概率分布获取随机机样本的程序，称之为随机数字生成器(random number generators)。例如，利用MINTAB统计软件，我们可从$N(0，1)$正态总体中得到25个随机数字，见表3-1的第1列，该表第2列是来自正态总体$N(2，4)$的一个随机样本[1]，当然，你可以利用有关统计软件生成许多个随机样本。

3.1.4 靴襻抽样法

与从某一概率分布中生成若干随机样本不同，我们可从这个概率分布中生成一个随机样本，然后，从这个样本再生成若干样本，这种“在样本中抽样”的方法，称为靴襻抽样法(bootstrap sampling)，靴襻抽样法的本质就是利用自身的资源(也即一个给定的样本)。

为了阐明这种方法，我们来看表3-2。表中第一列给出了来自标准正态分布$N(0,1)$的一个随机样本，共有25个观察值。现在想从这25个样本观察值中再抽取另外一个同样有25个观察值的随机样本，如何完成呢？从这25个观察值中随机抽取一个，记录其值,然后放回，然后再随机抽取一个，按此过程，我们可从原始样本中获得另外一个随机样本。这里关键之处是：每次从原始样本中抽取一个，然后将其放回，再抽取下一个。这个过程称为“重复抽样”。表中第2、3、4列给出了利用靴襻抽样法得到的三个新的随机样本。注意：在靴襻抽样法得到的2、3、4列样本中，原始样本中的一些数字可能只出现一次，或多次，也可能一次也不出现。例如，1列原始样本中的第6个数-2.161，它在样本2中出现3次，但原始样本中的最后一个数0.36822，在样本2中一次也没有出现。这就是靴襻抽样法的特点。

表3-2 来自$N(0，1)$的靴襻样本

$N(0,1)$	靴襻1	靴襻2	靴襻3
–1.562 43	–2.161 00	0.222 27	–2.161 00
2.317 59	2.315 79	–0.806 64	–0.806 64
–0.806 64	–2.080 59	–3.191 42	–0.512 32
–0.562 93	–2.161 00	–1.300 80	0.123 33
–3.191 42	–2.161 00	0.368 22	0.064 55
–2.161 00	2.317 59	–1.217 31	–0.790 13
–1.217 31	0.222 27	1.213 86	0.064 55
–0.940 25	–2.080 59	–0.806 64	–1.217 31
0.064 55	0.064 55	–0.940 25	–1.562 43
0.222 27	–1.269 99	–0.512 32	–0.790 13
0.491 46	0.123 33	1.213 86	–0.081 41
2.354 12	–3.191 42	2.317 59	–0.512 32
–0.512 32	–1.562 43	–2.080 59	2.354 12
–0.216 30	–0.465 15	0.491 46	–0.806 64
–2.080 59	–0.081 41	–1.300 80	–0.790 13
–1.300 80	–1.300 80	2.354 12	–0.790 13

1 MINTAB统计软件可从一给定均值和方差的正态总体中生成一随机样本。实际上，一旦从标准正态总体中得到一个随机样本，就很容易将它转化成不同均值和方差的正态分布。令 $Y=a+bZ$,其中Z是标准正态变量，Y 的概率分布如何呢？由于Y是Z的线性组合，因此，Y也服从正态分布(为什么？)。其特征数又是什么呢？

$E(Y)=E(a+bZ)=a+bE(Z)=a$　　因为，$E(Z)=0$(为什么？)

$\mathrm{Var}(Y)=\mathrm{var}(a+bY)=b^2\mathrm{var}(Z)=b^2$　　因为，$\mathrm{var}(Z)=1$(为什么？)

因此，$Y\sim N(a, b^2)$。于是，若将Z乘以a再加上b，就可得到一个来自于均值为a、方差为b^2的正态总体的随机样本。因而，若$a=2,b=2$，则$Y\sim N(2，4)$。

(续)

$N(0,1)$	靴襻1	靴襻2	靴襻3
0.12333	0.06455	−0.46515	0.12333
−1.26999	−0.80664	−0.21630	−2.08059
−0.46515	1.21386	−1.56243	−2.08059
1.21386	−2.08059	−0.08141	2.35412
−0.42875	−0.79013	−3.19142	−2.13041
−0.79013	−3.19142	−0.56293	−0.21630
−0.08141	−0.80664	−1.56243	−0.08141
−2.13041	−1.30080	−1.30080	−0.79013
0.36822	2.35412	0.36822	−0.08141

3.2 样本均值 $\bar{X}$的抽样分布或概率分布

在前一章，我们介绍了样本均值是总体均值的估计量 [见式(2-50)]，但由于样本均值是依据某一给定样本而定，因此其值也会因随机样本的不同而变化。也就是说，样本均值也可看作是随机变量，并且有其自己的概率分布函数，我们是否能够求出样本均值的概率分布函数呢？答案是肯定的，如果样本是随机抽取得到的话。在第 2章中我们直观地描述了随机抽样的含义，即总体中每一个体有同等机会被选入样本。但是，在统计学中，随机抽样(random sampling)有更特殊的意义。我们称 $X_1,X_2,\cdots,X_n$构成一容量为n的随机样本，如果所有的 X_S是从同一概率密度(即每个 X_i有相同的概率密度函数)中独立抽取得到的。因而称 X_S为独立同分布随机变量(independently and identically distributed random variables,i.i.d.random variables)。以后，随机样本都表示独立同分布随机样本。为了简便，有时用英文缩写 i.i.d.来表示独立同分布随机样本。

因此，如果$X_i \sim N(u，\sigma^2)$且每个X_i独立抽取得到，则称 $X_1,X_2,\cdots,X_n$是i.i.d.随机变量，正态概率密度函数是其共同的概率密度。对于这个定义，需注意两点：一是，样本中每一个 X有相同的概率密度函数，二是，样本中的每一个 X是独立抽取得到的。

定义了随机样本这一重要概念之后，现在讨论统计学中另一个非常重要的概念 —— 估计量(比如样本均值)的抽样分布或概率密度 [sampling,or probability,distribution of an estimator (e.g.,the sample mean)]。正确理解这一概念对于理解第 4章统计推断以及随后几章的内容十分重要，由于许多学生对这一概念的理解有些迷惑，因此，我们不妨用具体的实例来解释。

例3.6

正态分布的均值为10，方差为4，即$N(10，4)$。从这个正态总体中抽取 20个随机样本，每个样本包括 20个观察值。对抽取的每一个样本，得到其样本均值 $\bar{X}$，因而共有20个样本均值，见表3-3。

表3-3 来自$N(10，4)$的20个样本均值

样本均值($\bar{X}_i$)	
9.641	10.134
10.040	10.249
9.174	10.321
10.480	9.404
11.386	8.621
9.740	9.739
9.937	10.184
10.250	9.765
10.334	10.410
求和 =201.05	

$$\bar{\bar{X}} = \frac{201.05}{20} = 10.052$$

$$\text{var}(\bar{X}_i) = \frac{\sum (\bar{X}_i - \bar{\bar{X}})^2}{19} = 0.339$$

注：$\bar{\bar{X}} = \dfrac{\sum \bar{X}_i}{n}$

表3-4　20个样本均值的频率分布

样本均值范围	频数	频率
8.5～8.9	1	0.05
9.0～9.4	1	0.05
9.5～9.9	5	0.25
10.0～10.4	8	0.40
10.5～10.9	4	0.20
11.0～11.4	1	0.05
合计	20	1.00

表3-4给出了样本均值的频率分布，称为样本均值的经验抽样分布或概率分布。[1]图3-4的条线图描绘了样本均值的经验分布。

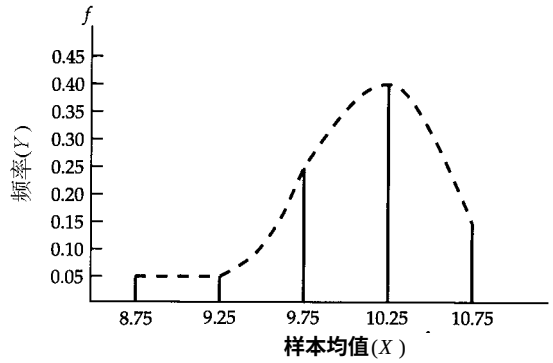


图3-4　来自$N(10，4)$总体的20个样本均值的分布

如果把图中各条线相连，就得到了频率多边形，其形状与正态分布的图形类似，如果我们列出更多这样的样本，那么频率多边形与正态分布的钟形曲线相类似吗？也就是说，样本均值的抽样分布确实服从正态分布吗？的确如此。

这里所依据的统计理论是：若$X_1,X_2,\cdots,X_n$是来自于均值为u，方差为σ^2的正态总体的一随机样本。则样本均值，$\overline{X}$也服从正态分布，其均值为u，方差为σ^2/n即

$$\overline{X} \sim N(u，\ \sigma^2/n) \tag{3-6}$$

换句话说，样本均值$\overline{X}$(u的估计量)的抽样(或概率)分布，同样服从正态分布。其均值与每一个X_i的均值相同，但方差等于X_i的方差($=\sigma^2$)除以样本容量n。[2]你会发现，当$n>1$时，样本均

1　一个估计量(比如样本均值)的抽样分布与任一随机变量的概率分布类似，只是此时的随机变量恰好是一个估计量，或者说，抽样分布就是当随机变量是一估计量时的概率分布，比如期望与方差，在后面的工作中你将遇到若干个这样的抽样分布。

2　证明：

因为，$\overline{X} = (X_1 + X_2 + \cdots + X_n)/n$

$$\begin{aligned} E(\overline{X}) &= [E(X_1) + E(X_2) + \cdots + E(X_n)] \\ &= [u + u + \cdots +]/n^2 \\ &= u \end{aligned}$$

$$\begin{aligned} \mathbf{Var}(\overline{X}) &= \mathrm{var}(X_1 + X_2 + \cdots + X_n)/n^2 \\ &= [\mathrm{var}(X_1) + \mathrm{var}(X_2) + \cdots \mathrm{var}(X_n)]/n^2 \\ &= (\sigma^2 + \sigma^2 + \cdots + \sigma^2)/n^2 \\ &= n\sigma^2/n \\ &= \sigma/n \end{aligned}$$

值的方差比任意一个X_i的方差都小。

回到刚才的例子中，则$\bar{X}$的期望值，$E(\bar{X})$，为10，方差为4/20=0.20。若取表3-3中20个样本的样本均值$\bar{\bar{X}}$，(称为总平均数)，它应该接近于$E(\bar{X}_i)$，样本方差应该约等于0.20。从表3-3中可知，$\bar{X}$=10.052, 约等于期望值10，方差var($\bar{X}_i$)为0.339，却与0.20相差甚远，为什么呢？

注意：表3-3中给出的数据仅仅来源于20个样本，前面我们强调过如果有更多的样本(每一样本有20个观察值)，那么，得到的结果将接近理论均值10和方差0.2。有这样的一个理论结果是令人鼓舞的。我们无须做诸如表3-3那样的抽样实验，那太消耗时间了。我们仅以一个来自于正态分布的随机样本为依据，就能说，样本均值的期望等于真实的均值u。在第4章中，将会看到，知道某一特定的估计量服从某一特定的概率分布将有助于建立从样本到总体之间的联系。这里略去计算过程，仅给出如下结果：

$$Z=\frac{\bar{X}-u}{\frac{\sigma}{\sqrt{n}}} \tag{3-7}$$

即Z是一个标准正态变量。因而，很容易从标准正态分布表中计算某一给定样本均值大于或小于某一给定的总体均值的概率。我们来看下面的一个例子：

例3.7

令X代表某一型号汽车每消耗一加伦汽油所行驶的距离(英里)。已知$X\sim N(20,4)$。则对于由一个有25辆汽车组成的随机样本，求：

(a)每消耗一加伦汽油所行驶的平均距离大于21英里的概率。

(b)每消耗一加伦汽油所行驶的平均距离小于18英里的概率。

(c)每消耗一加伦汽油所行驶的平均距离介于19和21英里之间的概率。

由于X服从均值为20，方差为4的正态分布，则$\bar{X}$也服从正态分布，其均值为20，方差为4/25。结果有，

$$Z=\frac{\bar{X}-20}{\sqrt{4/25}}=\frac{\bar{X}-20}{0.4}\sim N(0,1)$$

即Z服从标准正态分布，因此，现在要求：

$$(a)\ P(\bar{X}>21)=P\left(Z>\frac{21-20}{0.4}\right)=P(Z>2.5)=0.0062$$

$$(b)\ P(\bar{X}<18)=P\left(Z<\frac{18-20}{0.4}\right)=P(Z<5)\approx 0$$

$$(c)\ P(19\le\bar{X}\le 21)=P(-2.5\le Z\le 2.5)=0.9876$$

中心极限定理

刚才讲到从正态总体中抽样，其样本均值同样服从正态分布。但是如果从其他总体中抽样，情况又如何呢？这里有统计学中著名的中心极限定理(central limit theorem,CLT)：如果$X_1,X_2\cdots X_n$是来自(均值为u方差为σ^2的)任一总体的随机样本，随着样本容量无限增大，则其样本均值$\bar{X}$趋于正态分布，其均值为u，方差为σ^2/n。[1] 当然，如果X恰来自于正态分布，则无论样本容量为多大，其样本均值均服从正态分布，见图3-5，为了更清晰地描述，我们将连续型随型变量X在区间(a,b)上的均匀或矩形,概率密度函数(uniform,or rectangular,probability density function)定义为：

$$f(X)=\frac{1}{b-a}\qquad a<x<b \tag{3-8}$$
$$=0\qquad \text{其他}$$

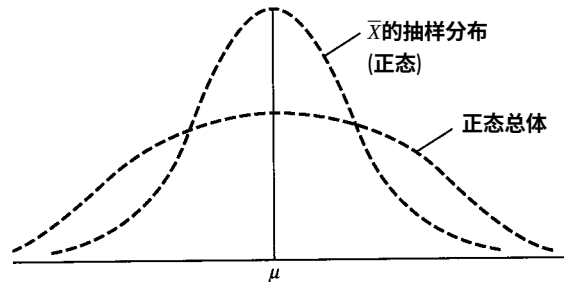
1 实际中，如果样本容量接近30，正态近似就已经很好了。在随后章节中将会看到，当样本容量大于30时，t分布近似正态分布。

图3-6给出均匀分布函数的图形。从图中可以看出，对于均匀分布的随机变量在其区间范围内任取一值的概率都相等，均为 $1/(b-a)$。若变量是离散型随机变量，则 X 在区间 (a,b) 取任一值的概率与 X 取所有值的概率都相等，事实上，我们在前面已经遇到过这种分布(见图 3-5)。掷一颗骰子，得到1至6中任何一个数字的概率均为 1/6。

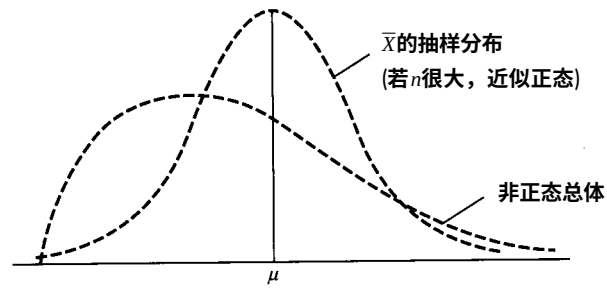
根据统计理论可以证明：均匀分布的均值和方差为：

$$E(X)=\frac{a+b}{2} \tag{3-9}$$

$$\text{var}(X)=\frac{(b-a)^2}{12} \tag{3-10}$$



a)　来自正态总体的样本



b)　来自非态总体的样本

图3-5　中心极限定理

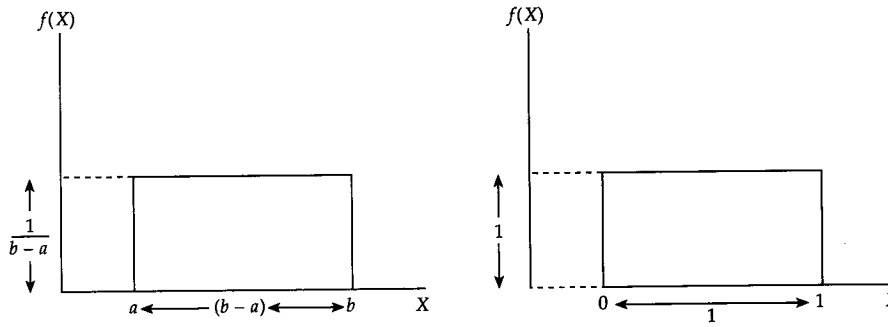


图3-6

a)均匀分布的概率密度函数 $U(a,b)$　b)标准均匀分布的概率密度函数 $U(0,1)$

为了阐明中心极限定理，从 $U(0,1)$ 总体中抽取20个随机样本，每个样本有 25个观察值。我们把在区间(0，1)上的均匀分布称为标准均匀分布(standard uniform distribution)，对每一个

样本，计算样本均值，共得到20个样本均值，见表3-5。

表3-5 来自标准均匀分布总体的20个样本均值

样本均值($\bar{X}$)	
0.54057	0.41617
0.41169	0.48899
0.43446	0.51373
0.51356	0.51264
0.46010	0.57346
0.47047	0.51848
0.46534	0.48087
0.42592	0.58560
0.57220	0.56843
0.47909	0.41683

$$\bar{\bar{X}}=\frac{\sum \bar{X}_i}{20}=0.492\,4 \qquad \text{样本方差}=\frac{\sum\left(\bar{X}_i-\bar{\bar{X}}\right)^2}{19}=0.003\,2$$

我们不是独立地看每一个样本均值，而是在一个更宽的区间范围内来考察。见图3-7，该图表明，虽然每一个X_i来自非正态分布总体，但样本均值$\bar{X}$都近似正态分布，当然，如果样本扩大到50、100或更多的话，则样本均值会更近似于正态分布，这也即是中心极限定理的含义。

在本例中，因为a=0, b=1，因此可以根据式(3-9)和式(3-10)验证，来自标准均匀分布总体的X_i的期望和方差分别为0.5和0.083 3。根据中心极限定理，$E(\bar{X})=0.5, \text{var}(\bar{X})=0.083\,3/25=0.0033$(注：$\text{var}(\bar{X})=\sigma^2/n$)。这个结果与表3-7进行的抽样试验相吻合吗？我们知道20个样本的总的均值($\bar{\bar{X}}$)为0.492 4，近似等于$E(\bar{X})=0.5$，方差为0.003 2，近似等于期望方差0.003 3。[1]正如刚才强调过的，如果样本容量足够大，则可以得到更好的近似值。

上述讨论的结论是：若样本容量足够大，则来自于任意分布总体的随机样本，其样本均值近似服从正态分布，见式(3-6)。[2]

上述进行的抽样试验又称为蒙特卡洛试验或蒙特卡洛模拟(Monte Carlo experiments or simulations)，它是研究统计模型的一个非常有用而又经济的方法；在经济计量学中，我们将大量地使用这一方法。

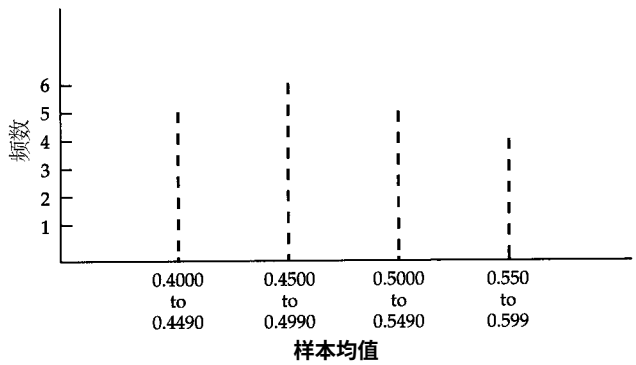


图3-7 来自标准均匀分布$U(0,1)$的20个样本均值的分布

1 样本方差的计算公式为：$\sum(\bar{X}_i-\bar{\bar{X}})^2/(n-1)$。

2 在统计还有一个更著名的理论，Linberg Fell理论：若X_s相互独立，且每一个X_s有自己的概率密度函数(并非所有的概率密度函数均不同)，则在一定条件下，当样本容易无限增大时样本均值仍然近似服从正态分布。

3.3 X^2分布

统计学中另一个常用的概率分布是χ^2分布[Chisquare(χ^2) distribution]。它与正态分布很相似。我们知道如果随机变量X服从均值为u，方差为σ^2的正态分布，即$X\sim N(u,\sigma^2)$，则随机变量$Z=(X-u)/\sigma$是标准正态变量，即$Z\sim N(0,1)$。统计理论证明：标准正态变量的平方服从自由度(degrees of freedom,d.f.)为1的χ^2分布，用符号表示为，

$$Z^2=\chi^2_{(1)} \tag{3-11}$$

其中χ^2的下标(1)表示自由度(d.f.)为1。正如均值、方差是正态分布的参数一样，自由度是χ^2分布的参数。在统计学中，自由度有不同的含义，但这里我们定义自由度是平方和中独立观察值的个数。[1] 在式(3-11)中，自由度仅为1，这是因为我们考虑的仅仅是一个标准正态变量的平方。

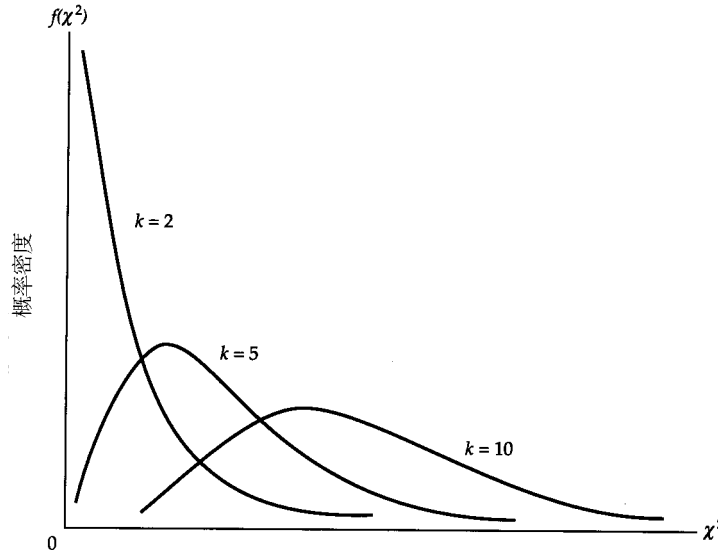


图3-8　χ^2变量的密度函数

现令$Z_1,Z_2,\cdots,Z_K$为K个独立的标准正态变量(即每一个变量均是均值为0，方差为1的正态变量)，对所有的变量Z_S平方，则它们的平方和服从自由度为K的χ^2分布，即

$$\sum Z_i^2 = Z_1^2 + Z_2^2 + \cdots + Z_k^2 \sim \chi^2_{(k)} \tag{3-12}$$

注意：这里的自由度为k，因为在式(3-12)的平方和中，有K个独立的观察值。

χ^2分布的几何图形见图3-8。

χ^2分布的性质

(1) 如图3-8示，与正态分布不同，χ^2分布只取正值(它是平方和的分布)且取值范围从0到无限大。

1　一般地，自由度的个数是指用于计算某个特征数(比如样本期望或样本方差)的独立观察值的个数；例如，随机变量X的样本方差定义为$S^2=\sum(X_i-\bar{X})^2/(n-1)$。在这种情况下，我们称其自由度为$(n-1)$，也就是说，如果我们用与计算样本方差相同的样本来计算样本均值时，将失去一个自由度，也即只有$n-1$个独立的观察值，我们举一个例子进一步说明，若X可取三个不同值：1、2、3，则样本均值为2。由于$\sum(X_i-\bar{X})=0$恒成立，所以，在差值(1−2)，(2−2)和(2−3)中只可任取2个，因为第三值必须满足条件$\sum(X_i-\bar{X})=0$。因此，在此情况下，虽然有三个观察值，但自由度仅为2。

(2) 如图3-8示，与正态分布不同，χ^2分布是斜分布，其偏度取决于自由度的大小，自由度越小，越向右偏，但随着自由度的增大，逐渐呈对称，接近正态分布。参见第 3.6节。

(3) χ^2分布的期望为k，方差为$2k$。k为χ^2的自由度，这是χ^2分布的一条特殊性质，即χ^2分布的方差是其均值的两倍。

(4) 若E_1、E_2分别为自由度为k_1,k_2的两个相互独立的χ^2变量，则其和(Z_1+Z_2)也是一个χ^2变量，其自由度为(k_1+k_2)。

附录A中表A-4给出了χ^2分布表。我们将在后面的章节中讨论 χ^2分布在回归分析中的作用，先介绍一下如何使用χ^2分布表。

例3.8

若自由度为30，求观察到的χ^2分布值大于13.18的概率？大于49的概率？大于50.89的概率？

从附录A中表A-4可查得所求概率分别为 0.995，0.95和0.01。因而，当自由度为30，χ^2分布值接近51的概率非常小，但对于同样的自由度，χ^2分布值较近14的概率就非常大，约为99.5%。

例3.9

随机样本来自方差为σ^2的正态总体，其样本容量为n，样本方差为S^2。可以证明：

$$(n-1)\left(\frac{S^2}{\sigma^2}\right) \sim \chi^2_{(n-1)}$$

即样本方差与总体方差的比值与自由度 $(n-1)$的积服从自由度为$(n-1)$的χ^2分布。假定来自正态总体 $(\sigma^2=8)$的一随机样本，样本容量为 20，若样本方差$S^2=16$，求得此样本方差的概率是多少？

我们将相应的数值代入上式,得$19\times(16/8)=38$。查附录A中表A-4得,当自由度为19时,如果真实方差$\sigma^2=8$,则获此χ^2分布值(为38)的概率为0.005。这里有一个疑问：这个特殊的随机样本是否来自于方差为8的正态总体呢？我们将在下一章给予详细地解答。

3.4 *t*分布

本书运用最广泛的一个概率分布是 *t*分布(*t* distribution)，*t*分布又称为学生*t*分布(Student's *t* distribution)，[1]它与正态分布也密切相关。

介绍*t*分布之前，先回忆一下,若$\bar{X}\sim N(u,\sigma^2/n)$，则变量$Z$服从标准正态分布：

$$Z=\frac{(\bar{X}-u)}{\sigma/\sqrt{n}}\sim N(0,1) \tag{3-13}$$

这里，假定u和σ已知。但是，现在假定仅知道 u及σ^2的估计量值S^2[见式(2-51)]。我们用样本标准差(S)代替总体标准差(σ)，得到一个新的变量

$$t=\frac{\bar{X}-u}{S/\sqrt{n}} \tag{3-14}$$

统计理论表明：变量t服从自由度为$(n-1)$的学生t分布。与χ^2分布类似，t分布也与参数自由度有关。这里，自由度为 $n-1$。注意：在用式(2-51)计算S^2之前，首先要计算样本均值 $\bar{X}$。但是

1 学生是W.S.Gosset的笔名，W.S.Gosset是一位统计学家，他于1908年发现了这一概率分布。

由于是用同样的样本来计算 $\bar{X}$，所以只有$(n-1)$个，而非n个独立的观察值来计算 S^2，也就是说，失去了一个自由度。

总之，从正态总体抽取随机样本，若该正态总体的均值为 u，但方差σ^2用其估计量S^2代替，则其样本均值服从t分布。t分布随机变量通常用符号 t_k表示，其中k表示自由度(为了避免为样本变量n混淆，我们一般地用下标 k表示自由度)附录 A中表A-2给出了对应不同自由度的 t分布值(t value)，稍后将介绍t分布表的使用。

t分布的性质

(1) t分布与正态分布类似，具有对称性，见图 3-9。

(2) t分布均值，与标准正态分布均值相同，为 0，但方差为$k/(k-2)$。因此，在求t分布的方差时定义自由度必须大于 2。

我们知道标准正态分布方差总为 1，这表明t分布方差总比标准正态分布方差大，见图 3-9，换句话说，t分布比正态分布略“胖”一些。

但是当k增大时，t分布的方差接近于标准正态分布方差值 1。因此，如果自由度 k=10，则t分布方差为10/8=1.25；如果自由度k=30，则其方差为30/28=1.07；如果自由度k=100，则其方差为100/98=1.02，仅比1略大一些 。因此我们可以得出结论：与 χ^2分布类似，随着自由度的逐渐增大时，t分布近似正态分布。但需注意的是，即使 k仅为30，t分布的方差已与标准正态分布方差相差不大。因此，对于 t分布，并不要求其样本容量很大，才能近似正态分布。

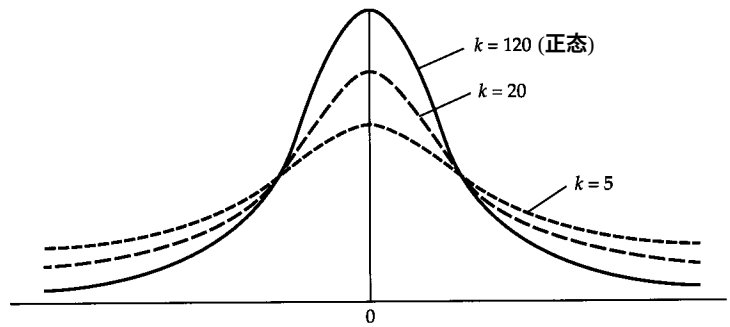


图3-9　不同自由度下的t分布

在说明如何使用t分布表之前，我们先来看一些具体的实例。

例3.10

再回到例3.2。在15天内，出售面包的平均数量为 14条(样本方差为4条)。假定真实的出售量为70条，求15内出售面包平均数量为74条的概率？

如果我们知道真实的标准差 σ，则可通过标准正态分布变量 Z来回答这个问题。但是，现在仅知道真实标准差的估计量 S，我们利用式(3-14)来计算t值，并用附录A中的表A-2回答这个问题：

$$t = \frac{74 - 70}{4/\sqrt{15}} = 3.873$$

注意：在本例中，自由度为 14=(15－1)(为什么？)。

当自由度为44时，查表A-2得，t值大于等于2.145的概率为0.025(2.5%)，t值大于等于2.624的概率为0.01(1%)，t值大于等于3.787的概率为0.001(0.1%)，因此，t值大于等于3.873的概率小于0.001。

例3.11

上例中其它条件保持不变，现假定15天内出售面包的平均数量为72条，求获此数量的概率？

接照上例步骤，读者容易求得此时的t值为1.936，再查表A-2得，当自由为14时，t值大于等于1.761的概率为0.05(5%)，大于等于2.145的概率为0.025(2.5%)；因此，t值大于等于1.936的概率位于2.5%与5%之间。

例3.12

现假定15天内，出售面包的平均数量为68条，样本方差为4，若真实的平均出售量为70条，求出售包平均数量为68条的概率？

将相应数字代入式(3-14)，得到此时t值为－1.936。由t分布的对称性可知，t值小于等于－1.936的概率与t值大于等于1.936的概率相同，则根据上例可知，所求概率在2.5%与5%之间。

例3.13

现求每天出售面包的平均数量在68条与72条之间的概率？

从例3.11和3.12可知，求平均出售量大于72或小于68的概率即为求t值大于1.936或小于－1.936的概率。[1]从前几例可知，这两个概率都介于0.025与0.05之间，因此，总的概率值介于0.05与0.10之间。在此例中，我们需计算$|t|>1.936$的概率，其中，$|t|$表示t的绝对值，即不考虑t值的符号。[2]

从上述各例可知，一旦根据式(3-14)计算出t值(自由度已知)，就可利用t分布表计算出给定某一t值的概率，t分布表在回归分析中的进一步应用，还将在后面作介绍。

例3.14

下面是1967～1990年间学生能力测试分数表：

	男	女
语言能力(平均分)	440.42	434.50
	(142.08)	(303.39)
数学(平均分)	500.0	453.67
	(46.61)	(83.88)

注：括号内的数字是方差值，数据来源于高校成绩统计，《纽约时报》，1990，8，28，表B-5。第5章的表5-5给出了实际数据。

由10位男生语言能力测试分数组成的随机样本。其样本均值和方差分别为440.60和137.60,若真实均值为440.42(全部1967～1990年期间),求样本均值为440.60的概率？

根据t分布知识，我们可回答这个问题。把相应的数值代入式(3-14)，得：

$$t=\frac{440.60-440.42}{\sqrt{\frac{137.60}{10}}}$$

在这个例子中,自由度为9(为什么？)。从表A-2中查得获此值的概率大于0.25(25%)。

1 这里需要仔细：比如－2.0小于－1.936，－2.3小于－2.0

2 例如，2的绝对值与－2的绝对值均为2。

3.5 F分布

F分布(F distribution)是经济计量学中又一种重要的概率分布。其定义如下：令随机样本$X_1,X_2,\cdots,X_m$来自均值为u_X，方差为σ_X^2的正态总体，其样本容量为m；随机样本$Y_1,Y_2,\cdots,Y_n$为来自均值为u_Y，方差为σ_Y^2为的正态总体，其样本容量为n。且这两个样本相互独立。假定想知道这两个正态总体是否同方差？即$\sigma_X^2=\sigma_Y^2$？由于我们不能直接观察两个总体的方差，但假定知道它们的估计量如下：

$$S_X^2=\sum\frac{(X_i-\bar{X})}{m-1} \tag{3-15}$$

$$S_Y^2=\sum\frac{(Y_i-\bar{Y})}{n-1} \tag{3-16}$$

现考虑比值：

$$F=\frac{S_X^2}{S_Y^2}=\frac{\sum(X_i-\bar{X})/(m-1)}{\sum(Y_i-\bar{Y})/(n-1)} \tag{3-17}$$[1]

如果两总体方差真实值确实相等，则根据式(3-17)计算出的F值将接近于1，但如果两总体方差真实值不相等，则F值不等于1；两总体方差相差越大，F值就越大。

统计理论表明：如果$\sigma_X^2=\sigma_Y^2$(即两总体同方差)，则比值F服从分子自由度为$(m-1)$，分母自由度为$(n-1)$[2]的F分布[3]。由于F分布常用于比较两总体的方差，因此，F[1]分布又称为方差比分布(variance ratio distribution)，通常用符号F_{k_1,k_2}表示，其中的双下标表明了分子与分母自由度(numerator and denominator d.f.)。[在此例中，$k_1=(m-1)$，$k_2=(n-1)$][4]

F分布的性质

(1) 与χ^2分布类似，F分布也是斜分布，向右偏，其取值范围也为0到无限大(见图3-10)。

(2) 与χ^2分布类似，当自由度k_1,k_2逐渐增大时，F分布近似正态分布。

(3) t分布变量的平方服从分子自由度为1，分母自由度为k的F分布，即

$$t_k^2=F_{1,k} \tag{3-18}$$

在第7章中我们将会看到这个性质的重要用处。

附录A中表A-3给出了F分布表。后面我们将看到F分布在回归分析中的特殊用途，首先介绍一下F分布表的使用。

---

1 通常，在计算F值时，将方差大的值放在分子上，这就是为什么F值总是大于或等于1。如果一个变量，比如说W，服从分子自由度为m，分母自由度为n的F分布，那么变量$(1/W)$同样也服从F分布，但此时，分子的自由度为n，分母的自由度为m。更特殊地

$$F_{(1-\alpha),m,n}=\frac{1}{F_{\alpha,m,n}}$$

其中α为显著水平，我们将在第4章中讨论。

2 注意：此时分子平方和，$\sum(X_i-\bar{X})^2$，的自由度为$(m-1)$，虽然有m个观察值。因为，在计算样本均值$\bar{X}$时失去了一个自由度。同样，分母平方和，$\sum(Y_i-\bar{Y})^2$，的自由度为$(n-1)$。因为，在计算样本均值$\bar{Y}$时失去了一个自由度。

3 更准确的，$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$服从$F$分布。但是，如果$\sigma_x^2=\sigma_y^2$，则$F$值由式(3-17)给出。

4 F分布有2个自由度。因为，F分布是两个独立变量与其各自自由度商的比值的分布。

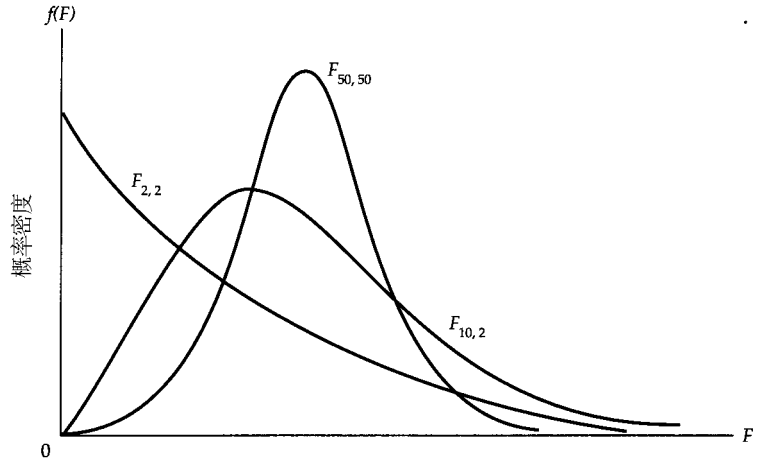


图 3-10

正如F分布与t分布存在特定的关系一样，F分布与χ^2分布也存在一定关系，

$$F_{(m,n)}=(\chi^2_m)/m \qquad \text{当} n\to\infty \tag{3-19}$$

即，χ^2变量与其自由度之比近似为分母自由度为 m，分子自由度很大(技术上说，无限大)的F变量；

因此，对于大容量样本，我们可用 χ^2分布来代替F分布；同样，也可用 F分布代替χ^2分布。例如，从F分布表中，可观察到，在 5%的显著水平下，$F_{3,120}=2.68$。从χ^2分布表中可观察到对自由度为3，在显著水平为5%下的χ^2临界值为7.814 73，因此，(7.814 73)/3＝2.604 9，与2.68相差不大。若现在我们考虑 $F_{3,\infty}$，从F分布表可查得，在 5%的显著水平下，F值为2.60，几乎与$(X^2_3/3)$相等。

同时，也可以将等式(3-19)写为：

$$mF_{m,n}=\chi^2_m \qquad \text{当} n\to\infty \tag{3-19a}$$

即，若分子自由度充分大(技术上说，无限大)，则 $F_{m,n}$值的m倍，等于自由度为 m的χ^2分布值。

例3.15

回到S.A.T一例中，(例3.14)，假定男、女生的语言能力的测试分数均服从正态分布，进一步假定它们的均值和方差来自于一个大容量总体。根据两个样本方差，能否认为两总体同方差？

因为男、女生的语言能力测试分数均为正态变量，因此，根据式(3-17)计算F值，

$$F=\frac{303.39}{142.08}=2.1353$$

其中，分子自由度为 23，分母自由度为 23。(注：在计算F值时，通常把方差值大的放在分子上)虽然附录A中表A-3并未给出自由度为 23时的F值，但当分子、分母自由度均为 24时，F值约为2.14的概率介于 1%～5%之间。如果我们认为这个概率值太小(详细的讨论见下章)，则可认为两总体方差不相等；即，男生的语言能力测试分数的方差与女生的不同。记住：如果两总体方差相等，则 F值为1。但若不相等，则F值将大于 1。

例3.16

两班做同样的经济计量学测试。其中，一个班级共有100名学生，另一班级共有150名学生，该老师从第一个班级随机抽取25个学生，从第二个班级随机抽取31个学生，观察得到两个班级学生考试平均分数的样本方差分别为100和132。假设学生考试平均分数这一随机变量服从正态分布，那么是否能够认为两班级分数平均值同方差。

因为这两个随机样本来自两个正态总体，并且相互独立。利用式(3-17)计算F值，

$$F=\frac{132}{100}=1.32$$

它服从自由度为30、24的F分布。查F分布表得当分子自由度为30、分母自由度为24时，F值大于等于1.31的概率为25%。如果我们认为这个概率相当大，就可得出结论：两总体同方差。

3.6 t分布、F分布、χ^2分布与正态分布的关系

在继续新的内容之前，需要指出：t分布、χ^2分布，F分布，及正态分布之间存在着密切的关系。因此，对它们之间的关系作一小结是必要的，因为，在后面的学习中我们也将会用得到这些关系。

(1) 若自由度充分大(至少为30)，则t分布近似标准正态分布。

(2) 若分母自由度充分大，F值的m倍(m为分子自由度)近似自由度为m的χ^2分布。

(3) 若$Z\sim N(0,1)$和χ^2_m相互独立，且χ^2分布的自由度为m，则

$$\frac{Z}{\sqrt{\chi^2/m}}=t_m \tag{3-20}$$

即，标准正态变量与χ^2变量与其自由度比值的平方根之比值服从自由度为m的t分布。在本书第6章中我们将介绍这个公式的使用。

(4) 根据式(3-18)，已经知道

$$t_k^2=F_{1,k}$$

(5) 若自由度充分大，则χ^2分布近似标准正态分布。

$$Z=\sqrt{2\chi^2}-\sqrt{2k-1}\sim N(0,1) \tag{3-21}$$

其中k为χ^2分布的自由度。举例见习题3.7。

3.7 小结

本章主要讨论了四种特殊的概率分布——正态分布、χ^2分布、t分布及F分布，并且概括了各个分布的一些重要性质，特别地介绍了这些概率分布的适用条件。在本书随后的章节中，您将会看到这四种概率分布在整个经济计量理论和实践中的重要作用。因此，对这四种概率分布的熟练掌握将是学习下面内容的基础，读者可以在今后的学习过程中，不时地翻回本章，仔细考虑这些分布的特殊之处。

习题

3.1 概念解释

(a) 自由度 (b) 估计量的抽样分布 (c) 标准差

3.2 若随机变量X～$N(8,16)$，判断正误：

(a) $P(X>12)=0.16$ (b) $P(12<X<14)=0.09$ (c) $P(X-u>2.5\sigma)=0.0062$

3.3 继续上题

(a) 求来自该总体的一随机样本的样本均值$\bar{X}$的概率分布； (b) (a)题的答案与样本容量有关吗？为什么？ (c) 假设样本容量为25，求样本均值为6的概率？

3.4 正态分布与t分布有什么区别？什么时候可用t分布？

3.5 假定随机变量服从t分布。

(a) 当自由度为20，求t值大于1.325的概率？ (b) 求t值小于1.325的概率？ (c) 求t值大于或小于1.325的概率？ (d) “求t的绝对值大于1.325的概率？”与(c)有区别吗？

3.6 判断正误：

若自由度充分大，则X分布，X^2分布，F分布都近似标准正态分布。

3.7 假定$k=50$，利用式(3-21)求解下列各题：

(a) 利用χ^2分布表求χ^2值大于80的概率？ (b) 利用正态近似求此概率？ (c) 现假定自由度为100，利用χ^2分布表计算此概率。

利用正态分布近似χ^2分布，你得出什么样的结论？

3.8 在统计学中，中心极限定理有何重要作用？

3.9 列举运用χ^2分布和F分布的例子。

3.10 变量X代表某一行业(由100个公司组成)的利润，假定X服从均值为150万美元，标准差为120 000美元的正态分布，计算：

(a) $P(X<100$万美元$)$ (b) $P(80$万美元$<X<130$万美元$)$

3.11 在习题3.10中，若有10%的公司的利润超过某一值，求此利润值？

3.12 假定经济计量学考试的平均分服从均值为75的正态分布，现有一随机样本(由学生总数的10%组成)，发现该样本的平均分数超过80，求此平均分数的标准差？

3.13 若一管牙膏的重量服从正态分布，其均值为6.5盎司，标准差为0.8盎司。生产每管牙膏的成本为50美分。若在质检中发现其中一管牙膏的重量低于6盎司，则需重新填充，重新填充每管牙膏的平均成本为20美分。另一方面，若牙膏的重量超过7盎司，则公司将每管损失5美分的利润，现检查1000支牙膏，

(a) 有多少管被发现重量少于6盎司？ (b) 在(a)情况下，重新填充而耗费的成本为多少？ (c) 有多少管牙膏重量多于7盎司？在此情况下，将损失多少利润？

3.14 若X～$N(10,3)$，Y～$N(15,8)$，且X和Y相互独立，求下列各式概率分布？

(a) $X+Y$ (b) $X-Y$ (c) $3X$ (d) $4X+5Y$

3.15 接上题，现假设X与Y正相关，相关系数为0.6，求上面各式的概率分布？

3.16 令X、Y代表两支股票的收益率，设X～$N(15,25)$，Y～$N(8,4)$，且X与Y的相关系数为-0.4。假设你持相同两支股票的比例相同。求股票收益率的概率分布？是仅投资于其中一支股票的收益大还是同时投资两种股票的收益大呢？为什么？

3.17 回到例3.14。随机样本包括10位女生的S.A.T数学分数，样本方差为85.21。已知真实的方差为85.21。求获此样本方差的概率？你用哪个分布回答这个问题？用此分布有哪些假定条件？

3.18 10位经济学家预测来年GDP的增长率。假定随机变量“预测”服从正态分布。

(a) 已知样本方差比总体方差大X个百分点的概率是0.10。求X。 (b) 若样本方差位于总体方差的X和Y之间的概率是0.95，求X和Y。

3.19 称10个谷箱的重量如下：(单位：盎司)

16.13 16.02 15.90 15.83 16.00
15.79 16.01 16.04 15.96 16.20

(a) 求样本均值和样本方差？ (b) 若真实的均值为16盎司，则得此样本均值的概率是多少？你用的是哪种概率分布？为什么？

3.20 两所大学微观经济学考试成绩如下：

$\overline{X}_1=75, S_1^2=9.0, n_1=50$

$\overline{X}_2=70, S_2^2=7.2, n_2=50$

其中，$\overline{X}_S$表示样本的平均分数；S_S^2表示样本方差；N_S表示样本容量。如何检验假设：两总体同方差？你将使用何种概率分布？用此分布有那些假定条件？

3.21 蒙特卡罗模拟。从自由度为10的t分布中抽取25个随机样本，每个样本包括25个观察值。计算每个样本的样本均值。求样本均值的抽样分布？并用图形加以说明。

3.22 重复习题3.21，但现在是自由度为8的χ^2分布。

3.23 重复习题3.21，但现在是分子自由度为10，分母自由度为15的F分布。

3.24 利用式(3-19)，在5%的显著水平下，比较$\chi^2_{(10)}$与$F_{10,10}$，$F_{10,20}$与$F_{10,60}$。你得出什么结论？

第 4 章

统计推断：估计与假设检验

具备了概率、随机变量、概率密度、概率密度的特征(比如期望值、方差、协方差、相关、条件期望值)的基础知识之后，我们将要讨论统计学中另一个重要内容：统计推断。一般地说，统计推断是根据来自总体的样本对总体(概率密度函数)的种种统计特征作出判断。

4.1 统计推断的含义

前面已经讲过，总体和样本在统计学中是两个非常重要的概念。总体是指我们所关注现象出现的可能结果的全体(例如，纽约的人口)。样本是总体的一个子集(例如，曼哈顿的人口，曼哈顿是纽约的五大城区之一)。更宽泛地说，统计推断研究的是总体与来自总体的样本之间的关系。我们通过一个具体的例子来说明统计推断的含义。

纽约股票交易市场(NYSE)共有1 758支股票(1990年9月4日)。假定某一天，我们从这 1 758支股票中随机选取50支，并计算这50支股票价格与收入比的平均值 —— 即P/E比值。(例如，一支股票的价格为50美元，估计年收益为5美元，则P/E为10；也就是说，股票以10倍的年收益出售。)[1]在统计推断中，提出这样一个问题：根据 50支股票的平均P/E值，能否说这个P/E值就是总体的1 758支股票的平均P/E值呢？换句话说，如果令X表示一支股票的P/E值，$\overline{X}$表示50支股票的平均P/E值$\left(=\sum_{i=1}^{50} X_i/50\right)$，我们能否得知总体的均值，$E(X)$呢？统计推断的实质就是从样本值$(\overline{X})$归纳出总体值$E(X)$的过程。

4.2 估计和假设检验：统计推断的两个孪生分支

从上面的讨论中可以看到统计推断是按下述过程进行的。首先有某一受关注的总体，比如说NYSE的1 750支股票，并且想要研究该总体某一方面的统计特征，比如说 P/E值。当然，并不是研究每一支股票的P/E值，我们只关心平均的P/E值。由于若要收集所有这 1 758支股票的P/E值，需要大量的计算，花费大量的时间和财力。因此，我们可以抽取一随机样本，比如 50支股票，求样本中每一支股票的 P/E值，然后再计算样本的平均 P/E值，$\overline{X}$。$\overline{X}$就称为总体平均

1 样本容量究竟取多大是另一类问题。统计学的另一个分支 —— 抽样理论深入地讨论了这类问题。但需要指出的是：如果样本确实是随机样本，就无须抽取一个容量非常大的样本。事实上，在美国总统选举投票中，其样本容量仅为1 000或1 500人，占美国总人口的比例非常小。

P/E[也即$E(X)$]的估计量，$E(X)$称为总体参数(回顾第2章：一个总体可由其参数来描述)。例如，均值和方差是描述正态分布的参数，估计量的某一取值称为估计值(例如，$\bar{X}$值为10)。因此，估计是统计推断的第一步。

得到参数的估计值后，接下来要判定估计值的“优度”，因为估计值很可能不等于真实的参数值：如果有两个或更多个随机样本，计算这些样本的均值$\bar{X}$，则得到的估计值很可能不相同。我们把不同样本估计值的差异称为抽样误差。[1]那么，是否存在一个判定估计量优劣的标准呢？在4.4节中我们将讨论判定估计量优劣的一些常用标准。

估计是统计推断的一个方面，假设检验则是统计推断的另一方面。在假设检验中，我们可对某一参数的假定值进行先验判断或预期，比如说，以往的经验或专家的意见告诉我们1 758支股票总体的平均P/E值为12，假定根据某一随机样本(样本容量为50)，计算出P/E的估计值为11。那么，11接近于假设值12吗？显然，两个数值并不相等。但是这里有一个重要问题：11与12显著不同吗？我们知道由于抽样的差异很可能导致样本估计值与总体真实值不同。“从统计上说”，或许11并不与12不同？在此情况下，我们能够拒绝假设：真实平均P/E值为12。但是如何作出判定呢？这就是假设检验的内容，我们将在4.5中详细讨论。

4.3 参数估计

在前面的章节中，我们讨论了几种理论概率密度。通常假定某一随机变量X服从某种概率密度，但并不知道其分布的参数值。举个例子，如果X服从正态分布，我们想知道其两个参数，均值$E(X)=u_X$，及方差σ^2。为了估计这些未知参数，一般的步骤是：假定有来自某一总体，样本容量为n的随机样本，根据样本估计总体的未知参数。因此，可将样本均值作为总体均值(或期望)的估计量，样本方差作为总体方差的估计量。这个过程称为估计问题。估计问题有两类：点估计(point estimation)和区间估计(interval estimation)。

假定随机变量X(P/E值)服从某一未知均值和方差的正态分布。但是，若有来自该正态总体的一随机样本(50个P/E值)，见表4-1。

如何根据这些样本数据计算总体的均值u_X(=$E(X)$)和方差σ^2呢？更具体的，假定现在仅仅关注u_X。[2]那么，如何计算u_X呢？根据表4-1的数据，50个P/E的样本均值为11.4，显然我们可以选择$\bar{X}$作为u_X的估计值。我们称这个单一数值为u_X的点估计值，称计算公式$\bar{X}=\sum_{i=1}^{50} X_i/50$为$u_X$的点估计量。注意：点估计量是一个随机变量，因为其值随样本的不同而不同。(回顾例3-6的抽样试验)那么，某一特殊的估计值(比如11.5)的可信度有多大呢？虽然$\bar{X}$可能是总体均值的“最好的”估计值，但是某个区间，比如8～10，更可能包括了总体均值。这也正是区间估计的基本思想。下面我们讨论区间估计。

区间估计的主要思想源于估计量抽样分布(或概率分布)的概念。我们知道，如果随机变量$X\sim N(u_X,\sigma^2)$，则

$$\bar{X}\sim N\left(u_X,\frac{\sigma^2}{n}\right) \tag{4-1}$$

或，

$$Z=\frac{(\bar{X}-u_X)}{\sigma/\sqrt{n}}\sim N(0,1) \tag{4-2}$$

1 注意：抽样误差不是任意的，它是由于随机样本而导致的。不同的样本包括的元素是不同的。因此，在根据样本分析时，抽样误差是不可避免的。

2 可根据类似的方法估计σ^2。

也即，样本均值的抽样分布也服从正态分布。[1]

前面我们已经讲过，通常σ^2未知，但可用其估计量$S^2=\sum(X_i-\bar{X})^2/(n-1)$代替，则有：

$$t=\frac{\bar{X}-u_X}{S/\sqrt{n}} \tag{4-3}$$

服从自由度为$(n-1)$的t分布。

在P/E一例中，共有50个样本观察值，因此，此时的自由度为49。查附录A中的t分布表(表A-2)，我们发现没有给出自由度为49时的t分布，但从其他书上可查到：

$$P(-2.009\,6\leq t\leq 2.009\,6)=0.95 \tag{4-4}$$

见图4-1，也即区间$(-2.009\,6,\ 2.009\,6)$包括t值的概率为95%。[2]$-2.009\,6$和$2.009\,6$称为临界t值(critical t values)。表明了在临界值区间内，位于t分布曲线下区域的比例(见图4-1)。把式(4-3)代到式(4-4)中，得到：

$$P(-2.009\,6\leq \frac{(\bar{X}-ux)}{S/\sqrt{n}}\leq 2.0096)=0.95 \tag{4-5}$$

整理得：

$$P\left(\bar{X}-\frac{2.0096S}{\sqrt{n}}\leq ux\leq \bar{X}+\frac{2.0096S}{\sqrt{n}}\right)=0.95 \tag{4-6}$$

式(4-6)即为真实u_x的一个区间估计量。

$$10.63\leq u_x\leq 12.36\text{(近似值)} \tag{4-7}$$

即为u_x的95%的置信区间，见图4-2。

表4-1 假设的样本(50支股票的P/E比值)

P/E	频数
6	2
7	2
8	5
9	6
10	5
11	7
12	5
13	4
14	3
15	4
16	6
18	1
合计：	50
均值=11.5	
样本方差=9.275 5	
样本标准差=3.045 6	
中位数=众数=11	

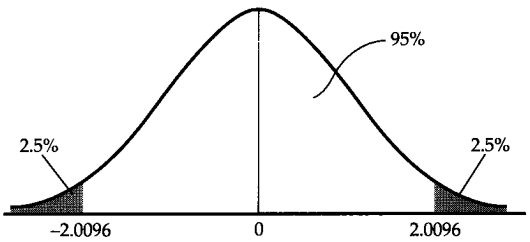


图4-1 t分布(d.f.=49)

在统计学中，称式(4-6)为未知的总体均值u_x的一个95%的置信区间。0.95称为置信系数(confidence coefficient)。一般的，式(4-6)表示随机区间$(\bar{X}\pm 2.009\,6S/\sqrt{n})$包括真实$u_x$的概率为0.95。$(\bar{X}-2.009\,6S/\sqrt{n})$称为区间的下限(lower limit)，$(\bar{X}+2.009\,6S/\sqrt{n})$称为区间的上限(upper limit)。

需要特别强调一点：式(4-6)给出的区间是随机的区间，因为它依赖于$\bar{X}$和$S/\sqrt{n}$的取值，而$\bar{X}$和$S/\sqrt{n}$的值随样本的不同而变化。虽然总体均值u_x是未知的，但是它取某一固定值，因而它是非随机的。所以，我们不能说u_x位于区间(4.6)的概率是0.95，只能说这个区间包括真实u_x的概率是0.95。简言之，区间是随机的，而参数u_x不是随机的。

回到P/E一例，根据表4-1知，$n=50$,$\bar{X}=11.5$，$S=3.045\,6$。把这些值代入式(4-6)，得：

式(4-7)表明，如果建立类似式(4-7)这样的置信区间100次，将有95次的区间包括真实的

1 如果X不服从正态分布，但若样本容量足够大，则根据中心极限定理，样本均值$\bar{X}$服从正态分布。

2 t值取决于自由度和显著水平，例如，同样的自由度，$P(-2.68\leq t\leq 2.68)=0.99$。

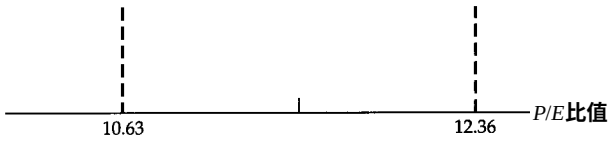


图4-2　总体平均的P/E的95%的置信区间

u_X。[1] 同时，在P/E一例中的下限和上限分别为10.63和12.36。

因此,与点估计(比如11.5)相比,区间估计提供了在某一置信度下(比如95%)真实参数的取值范围。

更一般地，假定随机变量X服从某一概率分布，若要对其参数进行估计，比如说，总体均值u_X。选取容量为n的随机样本，$X_1,X_2,\cdots,X_n$，并根据样本计算两个统计量(或估计量)L和U：

$$P(L \leq u_X \leq U)=1-\alpha \qquad 0<a<1 \tag{4-8}$$

也即，从L到U的随机区间包括真实u_X的概率为$(1-\alpha)$。L称为区间的下限，U称为区间的上限。这个区间称为u_X的置信区间。$(1-\alpha)$称为置信系数。如果$\alpha=0.05$，则置信系数为0.95，表示如果我们建立一个置信系数为95%的置信区间，那么重复建立这样的区间100次，预期有95次包括了真实的u_X。在实践中，通常将$(1-\alpha)$表示成百分比的形式，例如95%。在统计学中，α称为显著水平(level of significance)，或犯第一类错误的概率(probability of committing a type I error)，我们将在4.5节中详细讨论。

既然知道了置信区间是怎样建立的，那么置信区间有什么用处呢？在4.5节中将会看到，建立置信区间使假设检验——统计推断的一个分支——更为容易。

4.4　点估计量的性质

在P/E一例中，我们用样本均值作为u_X的点计量，同时还得到了u_X的区间估计量。但我们知道除了样本均值以外，样本中位数或样本众数同样可用作u_X的点估计量，为什么要选择样本均值为u_X的估计量呢？[2]从表4-1可知，样本中位数与样本众数均为11。

在实践中，样本均值是度量总体均值最广泛使用的统计量,因为它满足了下面的一些统计性质：

(1) **线性**(linearity)

(2) **无偏性**(unbiasedness)

(3) **有效性**(efficiency)

(4) **最优线性无偏估计量**(BLUE)

(5) **一致性**(consistency)

下面，我们来讨论这些性质。

4.4.1　线性

若估计量是样本观察值的线性函数，则称该估计量是线性估计量。显然，样本均值是一个线性估计量，因为$\overline{X}$是观察值X_S的线性函数：(注：X_S仅以一次幂的形式出现。)

1　这里需要特别注意。我们不能说每个特定区间包括真实的u_X的概率是0.95。因此，在古典的假设检验方法下，类似$P(10.63 \leq u_X \leq 12.36)=0.95$的表达式是不允许的。对区间式(4-7)正确的解释是：如果重复多次建立类似的区间，那么其中95%的区间包括真实均值；某一个特定的区间仅仅是区间估计量的一种实现形式。

2　中位数就是把总体PDF二等分的随机变量的取值，也即总体的所有取值一半大于该二等分值，一半小于该二等分值。根据样本求中位数，必须将样本按升序排列；中位数就是该顺序排列的中间值。例如，如果有观察值7，3，6，11，5，按升序排列为3，5，6，7，11。则中位数是6。众数也是描述随机变量常用值之一。例如，如果有观察值3，5，7，5，8，5，9。则众数为5，因为它出现的次数最多。

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

在统计学中，处理线性估计量比非线性估计量更为容易。

4.4.2 无偏性

如果总体某个参数有若干个估计量(即可用几种不同的方法估计参数)，并且如果平均而言，其中的一个或几个估计量与参数真实值相一致，就称这些估计量是无偏估计量。换种说法，如果重复使用某种方法，得到的估计量的均值与真实参数值一致，那么，这个估计量就是无偏估计量。更正规地，估计量，比如$\bar{X}$，称为无偏估计量，如果有：

$$E(\bar{X})=u_X \tag{4-9}$$

参见图4-3。如果情况不是这样，则称该估计量是有偏的估计量，比如图 4-3中的估计量X^*。

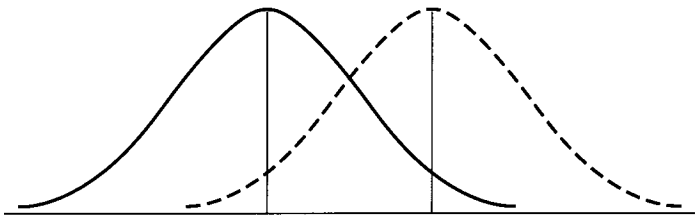


图4-3 总体均值u_X的无偏估计量$\bar{X}$和有偏估计量X^*

例4.1

若$X_i \sim N(u_X,\sigma^2)$，已知$\bar{X}$是来自该正态总体，样本容量为n的随机样本的样本均值，并且$E(\bar{X})=u_X$，$\mathrm{var}(\bar{X})=\sigma^2$。因此，样本均值$\bar{X}$是真实$u_X$的无偏估计量：如果从正态总体中重复抽取$n$个样本，并计算每个样本的样本均值$\bar{X}$，则平均而言，$\bar{X}$等于真实的$u_X$。但是需谨慎的是，我们不能仅通过一个样本 (比如表4-1这个样本)就认为计算的样本均值11.5一定就与真实的均值相一致。

例4.2

若$X_i \sim N(u_X,\sigma^2)$，假定从该正态总体中随机抽取容量为n的样本。X_{med}表示样本中位数。可以证明$E(X_{\mathrm{med}})=u_X$，即样本中位数也是真实均值的无偏估计量。需要注意：无偏性是一个重复抽样的性质；也就是说，如果抽取若干个容量为n的随机样本，计算每个样本中位数，则样本中位数的平均值接近u_X。我们不能保证一个中位数估计值(比如根据表4-1计算得到的样本中位数11)一定等于真实均值。

4.4.3 有效性

虽然无偏性是令人想望的性质，但却不是充分的。如果我们有两个或更多的无偏估计量？该作何选择呢？

假定随机变量X的取值构成一随机样本，样本的容量为n,并且每个$X \sim N(u_X,\sigma^2)$，令$\bar{X}$、X_{med}分别表示样本均值和样本中位数。已知：

$$\bar{X} \sim N(u_X,\sigma^2/n) \tag{4-10}$$

可以证明，若样本容量足够大，

$$X_{med} \sim N(u_X, (\pi/2)(\sigma^2/n)) \tag{4-11}$$

其中，π=3.142(近似值)。也即，对大样本而言，样本中位数也服从均值为 u_X 正态分布，但方差是样本均值 $\bar{X}$ 方差的(π/2)倍，从图4-4可以清楚地看出样本中位数的方差略大于样本均值的方差。实际上，如果求样本均值和中位数之比：

$$\frac{\text{var}(X_{med})}{\text{var}(\bar{X})} = \frac{\pi}{2}\frac{\sigma^2/n}{\sigma^2/n} = \frac{\pi}{2} = 1.571 \quad \text{(近似值)} \tag{4-12}$$

表明样本中位数的方差是样本均值方差的 1.571倍。

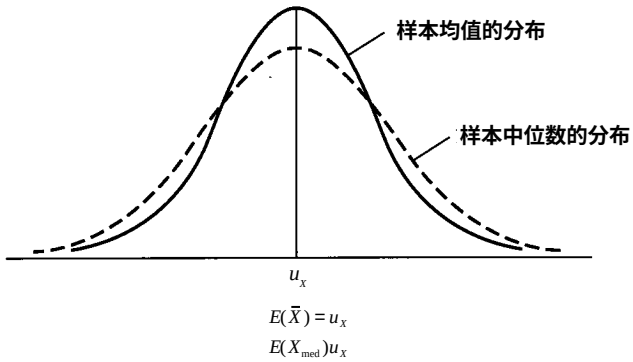


图4-4 有效估计量一例(样本均值)

根据图4-4及上述讨论，究竟选择哪一个估计量呢？一般地会选择 $\bar{X}$，因为虽然两个估计量都是无偏估计量，但 $\bar{X}$ 的方差比 X_{med} 的方差小。所以，在重复抽样中，如果用 $\bar{X}$ 估计 u_X 比用 X_{med} 更准确。简言之，$\bar{X}$ 提供了一个比 X_{med} 更为准确的总体均值的估计值。用统计语言，我们称 $\bar{X}$ 是有效估计量。用更正规的语言表达：若仅考虑惟一一个参数估计量，则方差最小的估计量是最好的或称为有效的估计量。

4.4.4 最优线性无偏估计量

在经济计量学中，最常遇到的一个统计性质就是最优线性无偏估计量(best linear unbiased estimator，BLUE)。如果一个估计量是线性的和无偏的，并且在参数的所有线性无偏估计量中，这个估计量的方差最小，则称这个估计量是最优线性无偏估计量。显然，这条性质包括了线性，无偏性及最小方差性，在第 6章和第7章中我们将会看到这些性质的重要作用。

4.4.5 一致性

为了解释一致性，假定 $X \sim N(u_X, \sigma^2)$，从该正态总体中抽取一容量为 n 的随机样本。现考虑 u_X 的两个估计量：

$$\bar{X} = \sum \frac{X_i}{n} \tag{4-13}$$

$$X^* = \sum \frac{X_i}{n+1} \tag{4-14}$$

第一个估计量就是常用的样本均值。我们知道

$$E(\bar{X}) = u_X$$

可以证明：

$$E(X^*) = \left(\frac{n}{n+1}\right)u_X \tag{4-15}$$[1]

既然$E(X^*)$不等于u_X，显然X^*是一个有偏的估计量。

但是，假定我们增大样本容量，情况又会怎样呢？估计量 $\bar{X}$ 与X^*的差别仅仅在于前者的分母为n而后者的分母为$n+1$。但是随着样本容量增大，发现两个估计量的差别不大。也就是说，随着样本容量的增加，X^*也将近似于真实的u_X，在统计学中，我们称这样的估计量为一致估计量。更正规的表述是：估计量(比如 X^*)称为一致估计量，如果随着样本容量的逐渐增大，该估计量接近参数的真实值。在后面的章节中我们将会看到，有些时候我们或许不能得到参数的无偏估计量，但却能得到一个一致估计量。[2]图4-5描绘了估计量的一致性。

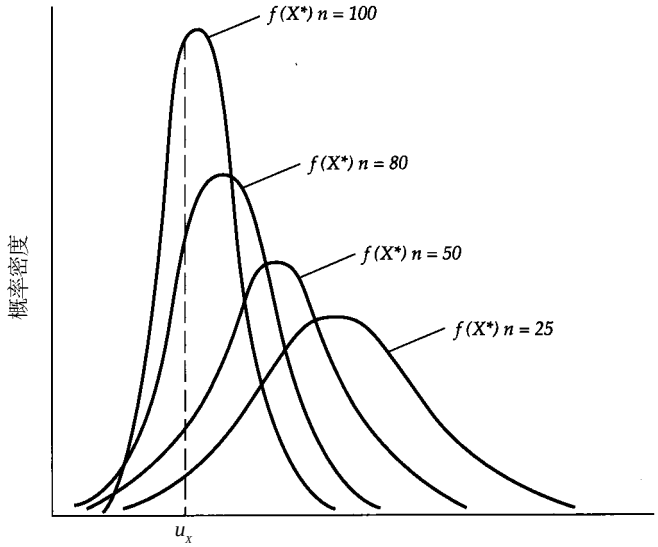


图4-5 随着样本容量的增大，总体均值估计量 X^*的变化

4.5 统计推断：假设检验

详细地研究了统计推断的分支之一——参数估计之后，现在讨论统计推断的另一重要分支，假设检验。虽然在前面已经讨论了假设检验的一般属性，但在这里，我们将更深入地讨论假设检验。

再来看表4－1的P/E一例。在4.3节中，根据50个P/E值组成的随机样本，我们对 u_X(真实的但未知的1 758支股票的平均P/E值)建立了一个95%的置信区间。现在改变策略，不是建立一个

1 很容易证明：

$$E(X^*) = \left(\frac{1}{n+1}\right)\{EX_1 + EX_2 + \cdots + EX_n\}$$
$$= \left(\frac{1}{n+1}\right)\{u_X + u_X + \cdots + u_X\}$$
$$= \left(\frac{n}{n+1}\right)u_X$$

这里我们用到：每一个X有相同的均值。

2 注意无偏估计量与一致估计量的不同之处。如果给定样本容量，从某正态分布总体抽取若干个随机样本，并根据随机样本估计总体参数，那么，无偏性要求平均而言，能够得到真实的参数值。另一方面，在建立一致估计量时，要看估计量随着样本容量的增加而产生的变化。如果样本容量足够大且估计量接近真实参数值，那么，这个估计量就是一致估计量。

置信区间，而是假设真实的 u_X取某一特定值，比如 u_X＝13。现在我们的任务就是去“检验”这个假设。[1]如何检验该假设呢？——也即是接受还是拒绝该假设？

用假设检验的语言，类似 u_X=13的假设称为零假设。通常用符号 H_0表示。因而，H_0：u_X=13。零假设通常与备择假设成对出现。用符号 H_1表示备择假设，备择假设有以下几种形式：

H_1：u_X>13 ，称为单边备择假设。

H_1：u_X<13 ，也称为单边备择假设。

H_1：u_X≠13，称为双边备择假设。即真实均值既不大于也不小于 13。[2]

为了检验零假设(与备择假设)，我们根据样本数据(比如，根据表 4-1得到的样本平均 P/E值 11.5)以及统计理论建立判定规则来判断样本信息是否支持零假设。如果样本信息支持零假设，我们就不拒绝 H_0，但如果不支持零假设，则拒绝 H_0，在后一种情况下，我们接受备择假设，H_1。

如何建立判定规则呢？有两个互补的方法：(1)置信区间法(2)显著性检验法。我们将通过 P/E一例来阐述这两种方法。这里，

$$H_0：u_X=13$$
$$H_1：u_X\neq 13 \quad \text{(双边假设)}$$

4.5.1 置信区间法

根据表4-1提供的样本数据计算出样本均值为 11.5。从4.3节讨论中，我们知道样本均值服从均值为 u_X，方差为 σ^2/n的正态分布。但是由于真实的方差是未知的，所以用样本方差来代替，在这种情况下，样本均值服从 t分布，见式(4-3)。根据 t分布，我们得到 u_X的一个95%的置信区间：

$$10.63\leq u_X\leq 12.36 \quad \text{(近似值)} \tag{4-7}$$

置信区间提供了在某一置信度下(比如 95%)真实的 u_X取值范围。因此，如果这个区间不包括零假设中的值，比如 u_X＝13，那么我们会拒绝零假设吗？答案是肯定的。我们以 95%的置信度拒绝该零假设。

从上面的讨论中，可以清楚地看到置信区间与假设检验密切相关。用假设检验的语言，不等式(4-7)描述的置信区间称为接受区域(acceptance region)(参见图4-2)，接受区域以外的称为零假设的临界区域(critical region)或拒绝区域(region of rejection)。接受区域的上界和下界称为临界值(critical values)。用这种语言表述为：如果参数值在零假设下位于接受区域内，则不拒绝零假设。但如果落在接受区域以外(也即落在拒绝区域内)，则拒绝零假设。在这个例子中，拒绝零假设 H_0：u_X=13，因为这个值落在临界区域，它比接受区域的上界 12.36大，也即这是一个小概率事件——不到2.5%。简言之，如果参数值超过上临界值或低于下临界值，那么就拒绝零假设。现在你就会清楚为什么接受区域的边界称为临界值。因为它们是接受或拒绝零假设的分界线。

4.5.2 第一类错误和第二类错误：一个偏离

在 P/E一例中，我们拒绝 H_0：u_X=13，因为样本均值 $\overline{X}$=11.5 看似与零假设不一致，这是否意味着表4－1给出的样本不是来自于均值为 13的正态总体呢？或许事实的确如此。别忘了不等式(4-7)的置信区间的置信度仅为 95%而并非100%。如果真的如此，那么拒绝 H_0：u_X=13，就可

1 假设就是“为了调查或讨论的目的，我们认为某件事是正确的”(webster’s)或是“基于某种原因之上的假定，或为了进一步调查而基于某些已知事实的一个出发点。”(牛津英汉词典)

2 零假设和备择假设有不同的表述方式。例如，零假设 H_0：$u_X\geq 13$,备择假设 H_1：$u_X<13$。

能犯错误。在这种情况下，我们说犯了第一类错误。也即弃真错误。同样的原因，假定零假设 H_0：u_X=12，在这种情况下，根据不等式(4-7)，我们应该不拒绝这个零假设。但是表4-1这个样本很可能不是来自均值为12的正态总体。因而，我们会犯第二类错误，也即取伪错误。

我们想尽可能减小这两种错误。但是，不幸的是，对于任一给定样本，我们不可能同时做到犯这两种错误的概率都很小。解决这一问题的古典方法是假定在实际中第一类错误比第二类错误更严重(由统计学家 Neyman和Pearson提出的)。因此，先固定犯第一类错误的概率在一个很低的水平上，比如说0.01或0.05，然后在考虑如何减小犯第二类错误的概率。[1]

犯第一类错误的概率通常用符号 α表示，称为显著水平，[2]犯第二类错误的概率用符号 β表示。则：

第一类错误=α=犯弃真错误的概率

第二类错误=β=犯取伪错误的概率

不犯第二类错误的概率(1-β)；也就是说，当 H_0为假时，拒绝 H_0，称为检验的功效(power of test)。

假设检验的标准或古典方法是：给定某一水平的 α，比如0.01或0.05，然后使检验的功效最大，也即使β最小。这个求解过程很复杂，有兴趣的同学可以参阅有关参考书。需要指出的是：在实际中，古典方法仅仅给出了α值，而没有过多考虑β值。

读者或许发现：前面讨论的置信系数(1-α)就是1减去“犯第一类错误的概率 α”，因此，95%的置信系数表示接受零假设犯第一类错误的概率至多为5%。简言之，5%的置信水平与95%的置信系数的意义相同。

我们用另一个例子进一步阐明假设检验的置信区间法。

例4.3

坛子里的花生的重量服从标准正态分布，但均值与标准差均是未知的，均值和标准差的度量单位为盎司*。随机选取20个坛子发现其样本均值和样本标准差分别为6.5盎司和2盎司。检验零假设：真实均值为7.5盎司；备择假设：真实均值不是7.5盎司。给定显著水平α=1%。

解答： 令X代表坛子中花生的重量，因此 $X\sim N(u_X,\ \sigma^2)$，两个参数u_X和σ^2均是未知的。由于真实方差是未知的，所以它服从自由度为19的t分布：

$$t=\frac{\bar{X}-u_X}{S/\sqrt{20}}\sim t_{19} \tag{4-16}$$

从附录A中表A-2的t分布表可知，自由度为19时，

$$P(-2.681\leq u_X\leq 2.681)=0.99 \tag{4-17}$$

根据式(4-16)可得：

$$P\left(\bar{X}-\frac{2.681S}{\sqrt{20}}\leq u_X\leq \bar{X}+\frac{2.681S}{\sqrt{20}}\right)=0.99 \tag{4-18}$$

将$\bar{X}$=6.5，S=2，n=20代入式(4-18)，我们得到了u_X的一个99%的置信区间。

$$5.22\leq u_X\leq 7.78 \quad \text{(近似值)} \tag{4-19}$$

由于式(4-19)的区间包括了零假设值7.5，因此，我们不拒绝零假设：真实的u_X=7.5。

---

1 这个过程听起来随意性很大，因为它没有仔细考虑两种错误的相对严重性。有关的详细讨论参见 Robert L.Winkler,*Introduction to Bayesian Inference and Decision*，Holt,Rinerhart and Winston，New York，1972，Chap.7。

2 α也称为(统计)检验的尺度。

* 1盎司=28.349 52g

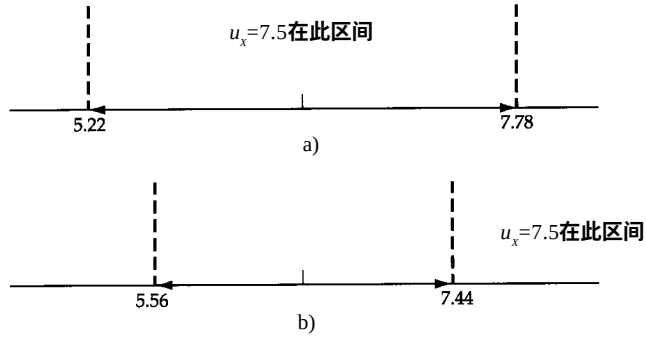


图4-6

a) u_x99%的置信区间 b) u_x的95%的置信区间(例4.3，例4.4)

例4.4

在例4.3中，若置信水平 a为5%；即决定冒更大的风险犯第一类错误。那么，情况如何呢？

解答：根据t分布表, 当a＝5%, 自由度为19时, t的临界值为-2.093和+2.093, 因为,

$$P(-2.093 \leq u_x \leq 2.093)=0.95 \qquad (4\text{-}20)$$

按照例4.3的步骤，读者能够求得：

$$5.56 \leq u_x \leq 7.44 \quad \text{(近似值)} \qquad (4\text{-}21)$$

从图4-7(b)中可以看出，这个区间不包括7.5，因此，拒绝零假设：u_x＝7.5。

式(4-19)与式(4-21)的不同之处在于后者的置信区间比前者略窄一些，得到这个结论我们并不应该感到奇怪，因为在建立不等式 (4-21)时，我们愿意冒较大的风险去犯第一类错误，即弃真错误。

例4.3，4.4表明，决定接受或拒绝零假设的关键在于自由度以及犯第一类错误的概率。

4.5.3 显著性检验

显著性检验(test of significance approach)是一种两者择一的假设检验，但它却是完备的。显著性检验是一种较为简洁的假设检验方法。我们仍通过P/E一例说明这种检验方法的一些基本要点。

根据式(4-3)可知：

$$t=\frac{\bar{X}-u_X}{S/\sqrt{n}}$$

服从自由度为$(n-1)$的t分布。在具体应用中，我们可知 $\bar{X}$，S，n。惟一未知的是u_x。但是如果设定u_x为某一值，正如在零假设下设定 u_x为某一给定值 ，则可以求出式(4-3)右边的值，因此，我们得到惟一一个t值。由于式(4-3)中的t服从自由度为$(n-1)$的t分布，因此，根据t分布很容易求得获此 t值的概率。如果 $\bar{X}$与u_x的差别不大(绝对值形式)，则根据式(4-3)可知，$|t|$也会很小($|t|$表示t的绝对值)。如果 $\bar{X}=u_x$，则t值为0，在此情况下，我们接受零假设。因此，随着$|t|$值偏离0，我们也将逐渐地趋向拒绝零假设：根据 t分布表，对给定自由度，$|t|$值越大，则获此$|t|$值的概率就越小。因而，随着$|t|$的逐渐增大，就越倾向于拒绝零假设。但在能拒绝零假设之前，最大的$|t|$值是多少呢？答案取决于置信水平 α，即犯第一类错误的概率，以及自由度。

这就是显著性检验方法的基本思想。这里的关键之处是检验统计量 —— t统计量 —— 以及在假定u_x为某一给定值下该 t统计量的概率分布。相应的，在本例中的检验是 t检验，因为我们用

的是t分布(对于t分布的详细讨论参见3.4节)。

在P/E例中，$\overline{X}$=11.5，S=3.045 6，n=50。令H_0：u_X=13，H_1：u_X≠13，因此有：

$$t=\frac{(11.5-13)}{3.045\ 6/\sqrt{50}}=-3.482\ 6$$

根据这个t值能否拒绝零假设？在没有设定置信水平以前，无法回答这个问题。假定 α=5%。由于备择假设是双边假设，因此，可以将犯第一类错误的风险均分在 t分布的两侧 —— 两个拒绝区域 —— 如果计算的t值位于任何一个拒绝区域，就能够拒绝零假设。

当自由度为49时，在5%的显著水平下，临界的t值为-2.009 6和2.009 6，见图4-1：获此t值小于或等于－2.009 6的概率为2.5%，获此t值大于或等于2.009 6的概率也为2.5%。

图4-1表明，本例中计算的t值约为－3.5，显然该t值位于t分布的左侧拒绝区域。因此，拒绝零假设，即真实的平均P/E值为13：如果该假设为真，则获此t值的概率不超过5% —— 犯第一类错误的概率。实际上，这个概率小于2.5%。(为什么？)

用显著性检验的语言，经常遇到下面两个术语：

(1) 检验(统计量)是统计显著的。

(2) 检验(统计量)是统计不显著的。

当我们说检验是统计显著的，一般是指能够拒绝零假设，即观察到的样本值与假设值不同的概率非常小，小于 α(犯第一类错误的概率)。同样的，当我们说检验是统计不显著的，是指不能拒绝零假设。在此情况下，观察到的样本值与假设值不同可能受抽样影响较大(即观察到的样本值与真实值不同的概率大于 α)。

当拒绝零假设时，我们就说是统计显著的，当不能拒绝零假设时，就说不是统计显著的。

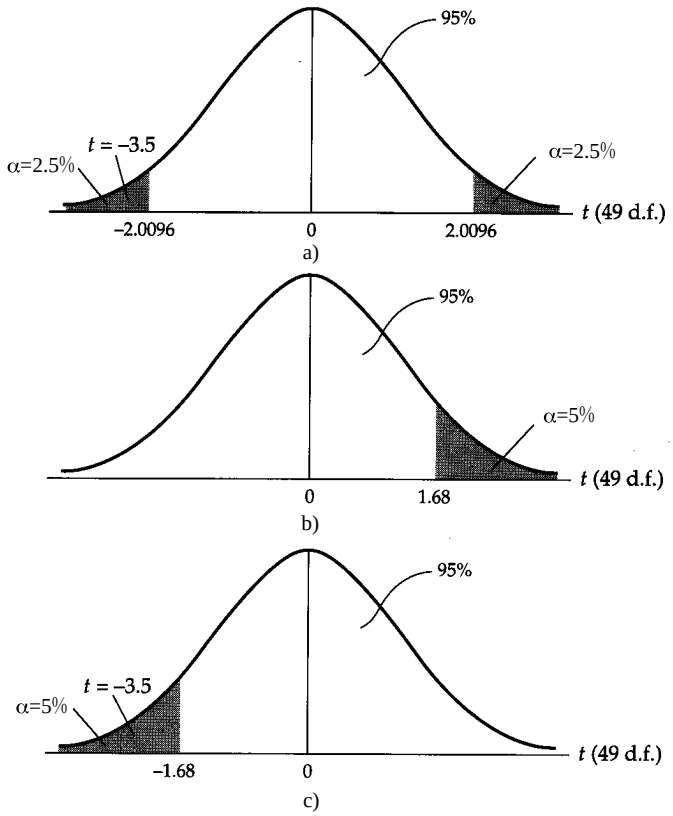


图4-7 t检验的显著性

a)双边检验 b)右侧单边检验 c)左侧单边检验

单边检验(one-tail test)**或双边检验**(two-tail test)

到目前为止，在我们考虑的例子中，备择假设都是双边的，或双侧的。因此，如果零假设是“平均P/E值等于13”，则备择假设下为“平均P/E不等于13”。在此情况下，如果检验统计量落在拒绝区域，那么就拒绝零假设。我们从图 4-7(a)中可清楚地看到。

然而，有些时候备择假设是单边假设。举个例子，在 P/E一例中，H_0：u_X=13，H_1：u_X<13。那么如何检验这个假设呢？

单边检验与前面讨论过的双边检验类似，只是在单边检验中，仅仅需要决定统计量单一的临界值,而不是两个临界值，见图 4-7。从图4-7可以看出，现在犯第一类错误的概率仅仅集中在该概率分布(t分布)的一侧。当自由度为 49,在5%的显著水平下，根据t分布表，得到单边的t临界值为1.676 6(右侧)或-1.676 6(左侧)，如图4-7示。在 P/E 一例中,计算的t值为-3.482 6(约等于-3.5)。由于t值在图4-7(c)的临界区域内,因此该t值是统计显著的。也即拒绝零假设：真实 P/E 值等于(或大于)13；发生 P/E 值等于(或大于)13的概率比选择的犯第一类错误的概率5%小的多。表4-2总结了单边检验和双边检验。

在实践中，是用置信区间法还是用显著性检验法，主要是取决于个人的选择与习惯。在置信区间方法中，我们对真实参数指定一个似乎合理的区间值,并查明参数假设值是落在该区间内还是落在区间外。如果落在区间内，我们就不拒绝零假设，但若落在区间外，则能够拒绝零假设。在显著检验方法中，我们不是对未知参数规定一个似乎合理的区间值，而是通过零假设给参数设定一个特殊值，再计算检验统计量，比如 t统计量，并求其抽样分布及该统计量取这个特殊值的概率。如果这个概率很低，比如说小于5%或1%，我们则拒绝零假设，如果概率值大于所选择的显著水平α，则不拒绝零假设。

表4-2　t检验小结

零H_0	假设备择假设H_1	临界区域，拒绝H_0，若
$u_X=u_0$	$u_X>u_0$	$t=\dfrac{\bar{X}-u_0}{S/\sqrt{n}}>t_{\alpha,\text{d.f.}}$
$u_X=u_0$	$u_X<u_0$	$t=\dfrac{\bar{X}-u_0}{S/\sqrt{n}}<-t_{\alpha,\text{d.f.}}$
$u_X=u_0$	$u_X\neq u_0$	$\lvert t\rvert=\dfrac{\bar{X}-u_0}{S/\sqrt{n}}>-t_{\alpha/2,\text{d.f.}}$

注：u_0表示在零假设下的u_X的某一取值。表中最后一列给出了 t临界值，t统计量的第一个下标代表了显著水平，第二个下标代表了自由度。

4.5.4　显著水平α的选择与P值

假设检验的古典方法的不足之处在于选择 α的任意性。虽然一般常用的 α值有1%、5%和10%，但是这些值并不是固定不变的。前面我们指出，只有在检查犯第一类错误和第二类错误后果的时候，才选择相应的α。在实践中，最好是用p值(即，概率值)，p值(p value)也称为统计量的精确置信水平。它可定义为拒绝零假设的最低置信水平。

我们用一个例子来说明。已知，当自由度为 20时，计算得到t值为3.552。根据附录A中t分布表，求出得此t值的概率值(p值)为0.01(单边的)或0.002(双边的)。也即在0.001(单边)或(0.002双边)水平下，t值是统计显著的。

在零假设：真实的P/E值为13下，我们得到t值为－3.5。

$$P(t<-3.5)=0.000\ 5$$

这就是t统计量的p值。我们说在0.005或0.05%显著水平下，这个t值是统计显著的。换句话说，

如果给定α=0.000 5，则在这个显著水平下，我们能够拒绝零假设：u_X=13。显然，这个p值是一个非常小的概率值，它比常用的α值小的多。事实上，在这个例子中，犯第一类错误的概率仅为0.0005(单边)。因此，比选择α=0.00 5更能拒绝零假设。一则规律：p值越小，越能拒绝零假设。

用p值的优点是避免了在选择显著水平 α(1%，5%，10%)时的任意性。举个例子，如果检验统计量的p值为0.135，如果接受α=13.5%，那么，这个p就是统计显著的(也即在此显著水平下拒绝零假设)。同样的，犯弃真错误的概率为 13.5%。

现在,许多统计软件都能计算各种统计量的 p值，而且在很多的研究报告中,也都给出了 p值。

4.5.5 χ^2显著性检验和F显著性检验

除了t检验外，在后面的章节中我们还将遇到建立在 χ^2分布和F分布之上的显著性检验。由于这些检验的基本原理相同，所以这里我们仅仅通过一两个例子来说明 χ^2检验和F检验的机理，在后面的章节中，我们还会看到这些检验进一步的应用。

1. χ^2检验

在第3章中(见例3.9)中已经证明了，如果 S^2表示来自方差为 σ^2的正态总体，容量为 n的随机样本的样本方差，那么，

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \tag{4-22}$$

即，样本方差与总体方差的比值与$(n-1)$的乘积服从自由度为$(n-1)$的χ^2分布。如果已知自由度n及样本方差 S^2，总体方差 σ^2未知，则根据 χ^2分布可对未知的总体方差 σ^2建立一个$(1-\alpha)$的置信区间，其机制与t检验类似。

但是如果在零假设 H_0下，给定 σ^2一个具体的值，则利用式(4-2)可以直接计算χ^2值，并根据χ^2分布表进行显著性检验。我们用一个例子来说明。

例4.5

假定随机样本来自正态总体，样本容量为 35，样本方差S^2=12。检验零假设：真实的方差为9；备择假设：真实的方差不等于 9。给定显著水平α=5%。

这里，H_0：σ^2=9，H_1：$\sigma^2 \neq 9$

解：把相应数字代入式(4-22)，得：χ^2=30(12/9)=40。根据附录A中的χ^2分布表，得此χ^2值的概率为0.10(自由为30)。由于这个概率值大于5%的显著水平，所以不能拒绝零假设：真实方差值为9。

表4-3总结了χ^2检验的各种类型的零假设及备择假设。

表4-3 χ^2检验小结

零H_0	假设备择假设H_1	临界区域，拒绝H_0，若
$\sigma_X^2=\sigma_0^2$	$\sigma_X^2>\sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{\alpha,(n-1)}$
$\sigma_X^2=\sigma_0^2$	$\sigma_X^2<\sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{(1-\alpha),(n-1)}$
$\sigma_X^2=\sigma_0^2$	$\sigma_X^2\neq\sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{\alpha/2,(n-1)}$
	或	$< \chi^2_{(1-\alpha/2),(n-1)}$

注：σ_0^2是σ_X^2在零假设下的取值。最后一列给出了临界的 χ^2值。χ^2第一个下标表示显著水平，第二个下标代表了自由度。

2. *F*检验

在第3章中我们讨论过，如果X、Y是来自两正态总体的随机样本，自由度分别为 m和n，则变量：

$$F=\frac{S_X^2}{S_Y^2}=\frac{\sum(X_i-\bar{X})^2/(m-1)}{\sum(Y_i-\bar{Y})/(n-1)}$$

服从自由度为$(m-1)$和$(n-1)$的F分布。这里假定两正态总体同方差。换句话说，H_0：$\sigma_X^2=\sigma_Y^2$。我们用式(3-17)的F检验检验该零假设。我们来看下面的一个例子。

例4.6

参考习题3.14男女学生S.A.T数学分数一例。男女学生 S.A.T分数的方差分别为46.61和83.88。其样本观察值分别为24和23。假设这些方差代表了来自于一更大总体的样本。检验假设：男女生S.A.T数学分数总体同方差。(α=1%)

解答：这里，F值为83.88/46.62=1.80(近似值)。该F值服从自由度均为23的F分布，根据F分布表，自由度为24(表中未给出自由度为23的值)，在1%的显著水平下，临界的F值为2.66。由于计算的F值为1.80，小于2.74，故它不是统计显著的。也即，在α=1%，不能拒绝两总体同方差。

表4-4总结了F检验。

表4-4　F统计量小结

零H_0	假设备择假设H_1	临界区域，拒绝H_0，若
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$\frac{S_1^2}{S_2^2}>F_{\alpha,ndf,ddf}$
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$\frac{S_1^2}{S_2^2}>F_{\alpha/2,ndf,ddf}$
	或	$<F_{(1-\alpha/2),ndf,dd}$

注：1. σ_1^2,σ_2^2是两个总体的方差。

2. S_1^2,S_2^2是两个样本的方差。

3. ndf,ddf分别代表了分子自由度和分母自由度。

4. 在计算F值时，将样本方差值较大的放在分子上。

5. 表中最后一列给出了临界的F值。在F下标中，第一个表示显著水平，第二个表示分子和分母自由度。

6. $F_{(1-\alpha)/2,ndf,dd}=\frac{1}{F_{\alpha/2,ddf,nd}}$

作为本章的概括，我们对统计检验的步骤做一总结：

第一步：表述零假设H_0和备择假设H_1

(例如，H_0：u_X=13，H_1：$u_X\neq13$)。

第二步：选择检验统计量(例如，$\bar{X}$)。

第三步：确定检验统计量的概率分布[例如，$\bar{X}\sim N(u_X,\sigma^2/n)$]。

第四步：选择显著水平α，即，犯第一类错误的概率，(但需记注有关p的讨论)。

第五步：选择置信区间法或显著检验方法。

置信区间法。根据检验统计量的概率分布，建立一个 $100(1-\alpha)\%$的置信区间，如果该区间

(也即接受区域)包括零假设值，则接受零假设，但若该区间不包括零假设值，则拒绝零假设。

显著检验法。根据这一检验方法，在零假设下，得到相关的统计量(例如t统计量)，并根据相应的概率分布(比如F分布或χ^2分布)计算该统计量取某一特殊值的概率。如果这一概率值小于事先选择的显著水平α值，则拒绝零假设，但若大于α，则接受零假设。如果你不想事先选择α，则可依据该统计量的p值。

需提醒注意的是，无论是用置信区间法或是显著性检验法，错误地拒绝或接受零假设的概率为$\alpha\%$。

本章讨论的假设检验方法将贯穿本书。

4.6 小结

根据样本信息对总体参数进行估计和借助样本信息进行假设检验是(古典)统计统断的两个主要分支。本章主要讨论了这两个方面的一些最基本的问题。

习题

4.1 区别概念

(a) 点估计量与区间估计量 (b) 零假设与备择假设 (c) 第一类错误与第二类错误 (d) 置信系数与显著水平 (e) 第二类错误与检验功效

4.2 解释概念

(a) 统计推断 (b) 抽样分布 (c) 接受区域 (d) 检验统计量 (e) 检验临界值 (d) 显著水平 (e) p值

4.3 解释概念

(a) 无偏估计量 (b) 最小方差估计量 (c) 最优的或最有效的估计量 (d) 线性估计量 (e) 最优线性无偏估计量(BLUE)

4.4 判断正误说明理由。

(a) 参数的估计量是随机变量，但参数本是非随机的或是固定的。 (b) 参数的无偏估计量，总是等于参数本身，(比如说u_x无偏估计量等于u_x)。 (c) 最小方差估计量不一定是无偏的。 (d) 有效估计量的方差最小。 (e) 估计量是最优线性无偏估计量，仅当抽样分布是正态分布时成立。 (f) 接受区域与置信区间是同一回事。 (g) 当拒绝可能为假的零假设时，才发生第一类错误。 (h) 当拒绝可能为真的零假设时，才发生第二类错误。 (i) 中心极限定理表明样本均值总是服从正态分布。 (k) 显著水平与p值是同一回事。

4.5 阐述置信区间法与显著性检验法的不同之处。

4.6 假定样本的自由度为40，得到t值为1.35。“由于p值介于显著水平5%与10%之间(单边)，所以它不是统计显著的。”你认为这句话对吗？为什么？

4.7 求下列Z值的临界值

(a) α=0.05(双边检验) (b) α=0.05(单边检验) (c) α=0.01(双边检验) (d) α=0.02(单边检验)

4.8 求下列t临界值

(a) n=4,α=0.05(双边检验) (b) n=4,α=0.05(单边检验) (c) n=14,α=0.01(双边检验) (d) n=14,α=0.01(单边检验) (e) n=60,α=0.05(双边检验) (f) n=200,α=0.05(双边检验)

4.9 假定一个国家的平均国民收入服从均值u=1 000美元，方差σ^2=10 000的正态分布。

(a) 求平均国民收入介于800美元～1 200美元之间的概率？ (b) 求平均国民收入超过1 200美元的概率？ (c) 求平均国民收入低于800美元的概率？ (e)“平均国民收入超过5 000美元的

概率几乎为零”为真吗？

4.10 继续习题4.9，假设现有一容量为1 000的随机样本，样本均值$\bar{X}$(平均收入)为900美元。(a) 若u＝1 000美元，求获此样本均值的概率？ (b) 根据样本均值，对u建立一个95%的置信区间，该区间是否包括u＝1 000。如果不包括，你得出什么结论？ (c) 用显著性检验方法，决定你是否拒绝假设u＝1 000美元，你用什么检验，为什么？

4.11 假设坛子中花生的数量服从均值为u，方差为σ^2的正态分布。对不同时期的质量检查表明：有5%的坛子花生重量低于6.5盎司，10%高于6.8盎司。

(a) 求u，σ^2？ (b) 求花生重量大于7盎司的坛子占总数的百分比？

4.12 下面的随机样本来自均值为u，方差为2的正态总体：

8，9，6，13，11，8，12，5，4，14

(a) 检验零假设H_0：u=5，备择假设：H_1：$u\neq5$ (α=5%) (b) 检验零假设H_0：u=5，备择假设：H_1：$u>5$ (α=5%) (c) 求(a)的p值？

4.13 假定有一来自正态总体(均值为u，标准差为σ)的随机样本，样本容量为10。经计算得其样本均值$\bar{X}$＝8，样本标准差S＝4，对总体均值建立一个95%的置信区间。你用哪种概率分布？为什么？

4.14 若$X\sim N(u=8,\sigma^2=36)$。根据25个样本观察值，知样本均值$\bar{X}$=7.5。(a) 求$\bar{X}$的抽样分布？ (b) 求$P(\bar{X}\leq7.5)$？ (c) 根据(b)的计算结果，这个样本值是否来自上述总体？

4.15 计算下列各p值？

(a)$t\geq1.72$, d.f=24 (b)$Z\geq2.9$ (c)$F\geq7.59$,d.f.= 3 ,20 (d)$\chi^2\geq19$，d.f.=30

注 如果不能得到准确的p值，求最优的近似值。

4.16 经计算得t统计量为0.68(d.f=30)。“由于即使在10%的显著水平下，该t值仍不是统计显著的，因此可安全地接受相应的假设。”你认为对吗？求获此统计量的p值？

4.17 若$X\sim N(u,\sigma^2)$，现有来自总体的随机样本X_1,X_2,X_3。考虑下面的u的估计量：

$$u_1=\frac{X_1+X_2+X_3}{3},u_2=\frac{X_1}{6}+\frac{X_2}{3}+\frac{X_3}{2}$$

(a) u_1是u的无偏估计量吗？u_2呢？ (b)若u_1，u_2均为u的无偏估计量，你将选择哪一个估计量？ (提示：比较两估计量的方差)。

4.18 参见第3章的习题3.10。假定现有10个公司组成的随机样本，其平均利润为900 000美元，标准差为100 000美元。(a) 对整个行业的真实平均利润建立95%的置信区间？ (b) 你将用哪种分布？为什么？

4.19 参考第3章例3.9。

(a)求σ^2的置信区间？(α = 5%) (b)检验假设：真实的σ^2=8.2。

4.20 16辆汽车先使用标准燃料，再用石油燃料(汽油中混有甲烷)。一氧化氮(NO_x)的排放量如下：

燃料种类	平均NO_x的排放量	NO_x排放量的标准差
标准燃料	1.075	0.579 6
汽油	1.159	0.613 4

资料来源：Michael O.Finkelstein ，Bruce Levin，*Statistics* for *Lawers*，Springer-Verlag,New York，1990，p.230.

(a) 如何检验假设：两总体的标准差相同？

(b) 你用的是什么检验？用此检验有哪些基本假定？

第二部分

线性回归模型

第二部分共包括五章内容，旨在向对读者介绍经济计量学的基础工具——线性回归模型。

第5章通过最简单的线性回归模型——双变量模型探讨了线性回归的基本思想；区别了总体回归模型与样本回归模型（用后者估计前者），并介绍了常用估计方法——最小二乘法。

第6章讨论了假设检验。在统计学中，假设检验用来探求回归模型中参数的估计值是否与参数的假设值相一致。我们对古典线性回归模型（CLRM）进行假设检验。本章讨论了为什么要用古典线性回归模型，并且指出古典线性回归模型是研究其他回归模型的起点。在本书的第三部分中，我们将重新考虑古典线性回归模型的基本假设，看看当一个或几个假设不满足时，古典回归模型会有什么变化。

第7章，我们把前面两章建立起的双变量回归模型的思想推广到多元回归模型之中，即模型中的解释变量不止一个。虽然，在许多方面，多元回归模型是双变量模型的推广，但在对模型系数的解释以及假设检验的过程等方面，还是存在着一些差异。

线性回归模型，无论它是双变量模型还是多变量模型，我们仅要求模型的参数是线性的，而进入模型的变量本身并不要求是线性的。在第8章中就讨论了参数是线性（或能够转化成线性）而变量并非线性的模型。我们将通过若干具体的例子，指出什么时候以及如何使用这些模型。

有些时候，进入模型的解释变量是定性的，比如性别、颜色、宗教信仰等。通过第9章的讨论，您将了解如何使这些变量变为可以“度量”的变量，并且通过考虑这些变量的影响，知道它们是如何丰富线性回归模型的。

整个第二部分，我们都努力将理论用于实践。虽然，“用户友好”回归软件的运用使你无须知道太多的理论就会可以对回归模型进行估计，但是，别忘了“一知半解是危险的”。因此，即使理论很枯燥，它对理解和解释回归结果也是绝对必要的，此外，我们省略了所有的数学推导，这将会使理论变得简单一些。

第5章

线性回归的基本思想：双变量模型

第一章我们讲到在对经济现象(例如需求法则)建立经济计量模型时，经济计量学大量地使用了回归分析这一统计技术，本章和下章将通过最简单的线性模型“—”双变量模型来介绍回归分析的基本思想，在随后的几章中，我们将讨论对双变量模型的修正及推广。

5.1 回归的含义

回归分析是用来研究一个变量(称之为被解释变量(explained variable)或应变量(dependent variable))与另一个或多个变量(称为解释变量(explanatory variable)或自变量(independent variable))之间的关系。

我们也许对研究商品的需求量与该商品的价格、消费者的收入以及其他同类商品的价格之间的关系感兴趣；也许对研究产品的销量(比如，汽车)与用于产品宣传的广告费之间的关系感兴趣；也许对研究国防开支与国民生产总值(GNP)之间的关系感兴趣；在上述各例中，或许存在某一基本理论，它规定了我们为什么期望一个变量是非独立的或说它与其他一个或几个变量有关。在第一例中，需求法则就提供了这样一个理论基础——一种产品的需求量依赖于该产品的价格以及其他几个变量。

为了统一符号，从现在起，我们用 Y代表应变量，X代表自变量或解释变量。如果有多个解释变量，我们将用适当的下标，表示各个不同的 X。(例如，X_1，X_2，X_3等等)。

时刻记住第1章中的告诫是非常重要的：虽然，回归分析是用来处理一个应变量与另一个或多个自变量之间的关系，但它并不一定表明因果关系的存在；也就是说，它并不意味着自变量是原因，而应变量是结果。两个变量是否存在因果关系，必须以(经济)理论为判定基础，正如前面讲到的需求法则，它表明：当所有其他变量保持不变时，一种商品的需求量依赖于(反向)该商品的价格。这里，微观经济理论暗示了价格是原因，而需求量是结果。总之，回归并不意味着存在因果关系，因果关系的判定或推断必须依据经过实践检验的相关理论。

回归分析可以用来：

(1) 通过已知变量的值来估计应变量的均值。

(2) 对独立性进行假设检验——根据经济理论建立适当的假设。例如，对前面提到的需求函数，你可以检验假设：需求的价格弹性为－1.0；即需求曲线具有单一的价格弹性。也就是说，在其他影响需求的因素保持不变的情况下，如果商品的价格上涨1%，平均而言，商品的需求量将减少1%。

(3) 通过自变量的值，对应变量的均值进行预测。因此，对于前面讨论的 S.A.T一例，我们可以通过学生的数学成绩来预测其语言表达能力的平均分数。(见表 5-5)

(4) 上述多个目标的综合。

5.2 总体回归函数：一个假设的例子

为了说明上面提到的回归分析的用途，我们以需求法则为例。回忆一下，需求法则是说当影响需求量的其他变量保持不变时(经济学中著名的“其他条件不变”)，商品的需求量与其价格呈反方向变动关系。这些其他的变量包括消费者的收入、偏好、同类商品的价格、以及互补商品的价格等等。

下面我们来考虑对《widget》教科书的需求。假设有一小镇，镇上共有 55个人，表5-1给出了这些消费者对《widget》一书的需求量。

表5-1 对widget需求表

价格 (X)	需求量 (Y)	消费者数量	平均需求量
(1)	(2)	(3)	(4)
1	45, 46, 47, 48, 49, 50, 51	7	48
2	44, 45, 46, 47, 48	5	46
3	40, 42, 44, 46, 48	5	44
4	35, 38, 42, 44, 46, 47	6	42
5	36, 39, 40, 42, 43	5	40
6	32, 35, 37, 38, 39, 42, 43	7	38
7	32, 34, 36, 38, 40	5	36
8	31, 32, 33, 34, 35, 36, 37	7	34
9	28, 30, 32, 34, 36	5	32
10	29, 30, 31	3	30
	总计	55	

从表5-1中可以看出：当每本书的价格为 1美元时，有7个消费者愿意购买此书，他们对读书的需求量分别在45到51本之间，7个消费者的平均需求量为 48本，这一数值是通过对 7个人的需求量求算术平均数而得到。类似地，当每本书的价格为 7美元时，有5个消费者愿意购买此书，他们对读书的需求量分别在 32到40之间，在此价格下的平均需求量为 36个单位。表中其他数值可类似地解释。

以需求量(Y)为纵轴，以价格(X)为横轴，对表中的数据作散点图(见图 5-1)。从散点图可以看出：任一给定X，有若干个Y与之对应，例如，当$X=1$时，有7个值与他对应，当$X=4$时，相应地有6个Y值，等等。换句话说，一个X与一个Y总体(与X有关的Y值的全体)有关。(回顾第 2章中对总体的定义)

这个散点图能说明什么问题呢？给人的第一印象是，当 X增加时，Y一般减少。反之，当X减少时，Y一般增加，如果把注意力集中在图中圈起的点(即与 X_s相应的Y的总体均值)，这种趋势就显得更加明显。这些在图中圈起的点给出了预期均值(expected mean)或者总体均值(population mean)或者与X相对应的Y的总体平均值(population average value)。我们可以认为，平均的Y值随X线性减少(即是一条直线)。我们称该直线为总体回归直线(Population Regression Line，PRL)。总体回归直线给出了与每个自变量(这里是 X)相对应的应变量的均值，在我们的这个例子中，可以看到，当 X为2美元时，平均的需求量为 46个单位，虽然从表 5-1中得知在此价格下的需求量在44到48单位之间不等。简言之，总体回归直线告诉我们对应于每一个 X值(或任一个自变量)相应的Y(或任何一个应变量)的均值。

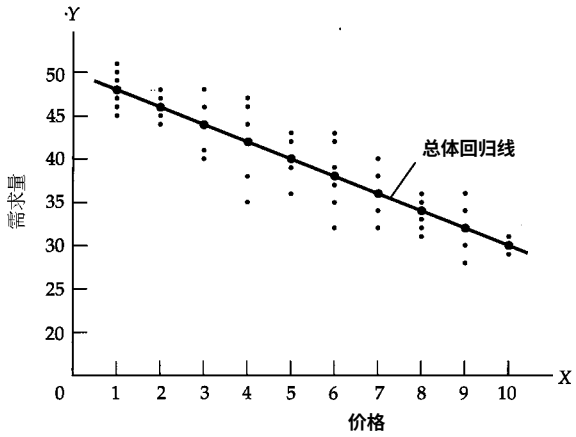


图 5-1

图5-1画出的总体回归直线是线性的，我们也可用函数的形式来表示：

$$E(Y \mid X_i)=B_1+B_2X_i \tag{5-1}$$

这是直线的数学表达式，在式(5-1)中，$E(Y \mid X_i)$表示给定X值相应的(或条件的)Y的均值，在第2章中，我们讲过这是Y的条件期望(conditional expectation)或条件均值(conditional expected value)。下标i代表第i个子总体。因而，$E(Y \mid X_i=2)=46$，即在第2个子总体中，Y的期望值或均值为46。

表5-1中最后一列给出了Y值的条件期望。需提醒注意的是，$E(Y \mid X_i)$是X_i的函数(在此例中是线性函数)。这意味着Y依赖于X_i，一般称之为Y对X的回归。回归可简单地定义为在给定X值的条件下Y值分布的均值。换句话说，总体回归直线经过Y的条件期望值。式(5-1)是总体回归函数(Population Regression Function, PRF)的数学形式。在本例中，总体回归函数是线性函数(线性更专业的定义将在5.6节中给出)。

在式(5-1)中，B_1，B_2为参数(parameters)，也称为回归系数(regression coefficients)。B_1又称为截距(intercept)，B_2又称为斜率(slope)。斜率度量了X每变动一单位，Y的均值的变化率。例如，如果斜率为－2，那么，当价格每提高1美元，Y的(期望)均值将减少2个单位；即，平均而言，需求量将下降2个单位。B_1是当X为0时的Y的均值。我们将在后面的章节中对截距作更多的解释。

如何求斜率和截距的估计值(或数值)呢？我们将在5.8节中探讨这个问题。

在继续下面内容之前，关于术语还有一点需提醒大家注意。在第1章中已经强调过我们所关注的是：在给定自变量的条件下，应变量的行为。因此，回归分析是条件回归分析(conditional regression analysis)。我们无需每时每刻都加上“条件”二字。所以，往后，表达式$EY(Y \mid X_i)$将简写为$E(Y)$，读者必须清楚后者是前者的简略写法。当然，在容易混淆的地方，我们仍将沿用完整的符号。

5.3 总体回归函数误差的设定

我们刚才讨论过，总体回归函数给出了对应于每一个自变量的应变量的平均值。让我们再看一下表5-1。当X＝1美元时，相应的Y的均值为48。但是，在这个价格水平下，从7个人中随机抽取一个，则他的需求量并非一定等于48。更具体地说，从这一组中取最后一人，则他的需求量为51个单位，位于均值之上。同样从这一组中选取第一个人，则其需求量为45个单位，在平均值之下。我们如何解释在某一价格下，个人的需求量呢？最好的方法是说个人的需求量等

于在平均的需求量上加上或减去某一数量。用数学公式表示为：

$$Y_i=B_1+B_2X_i+u_i \tag{5-2}$$

其中，u_i表示随机误差项(stochastic, random error term)或简称为误差项。[1]

我们在第1章中已经遇到过这一概念。误差项是一随机变量，因为其值不能先验地知道。在第2章中讲到，我们通常用概率分布(例如，正态分布式 t分布)来描述随机变量。

如何解释式(5-2)呢？我们可以说在某一价格水平上，个人的需求量，比如第 i个人的需求量，可看做两部分之和：(1)$(B_1+B_2X_i)$，第i个子总体的平均需求量，也即在此价格水平下总体回归直线上相对应的点。这一部分称为系统的或决定的部分。(2)u_i，称为非系统的或随机的部分(也就是说，他由价格以外的因素所决定)。

为了更清晰地理解，我们来看图5-2(是根据表5-1中的数据得到的)。

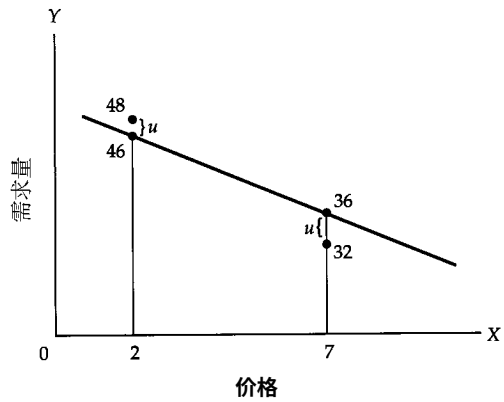


图　5-2

从图中可以看出，当 $X=2$美元时，某一消费者的需求量为 48个单位，但在此价格水平下的平均需求量为46个单位。因而，该消费者的需求量超过系统部分(即该子总体的均值) 2个单位。也就是说，u_i为2个单位；当 $X=7$美元时，随机抽取的一消费者的需求量为 32个单位，但在此价格水平下的平均需求量为 36个单位。因而，该消费者的需求量低于系统部分 4个单位，其u_i为－4个单位。

式(5-2)称为随机总体回归方程(stochastic PRF)，式(5-1)称为决定的(或非随机的)总体回归方程(deterministic or nonstochastic PRF)，后者表示对于具体价格，各个 Y值的均值是多少。而前者告诉我们由于误差项u_i的存在，个人需求量在其均值附近是如何变化的。

5.4　随机误差项的性质

(1) 随机误差项可能代表了模型中并未包括的变量的影响。例如在 widget一例中，它可能代表了诸如消费者收入，同类竞争产品的价格等因素的影响。

(2) 即使模型中包括了所有决定需求量的有关变量，需求量的内在随机性也一定会发生，这是我们做何种努力都无法解释的。毕竟，即使人类行为是理性的，也不可以完全可预测的。正因为如此，u_i或许反映了人类行为中的一些内在随机性。

(3) u_i也可以代表测量误差，例如，对需求量 Y的样本观察值，由于在数据统计时的四舍五

1　英文“ stochastic ”来自于希腊词“ stokhos ”，意思是“ 公牛的眼睛 ”。向标盘投掷飞标就是一个随机过程，即该过程可中标也可未中。在统计学中，这一词暗合随机变量的结果由随机试验决定。

入，都不可避免地会产生误差。

(4) Occam的剃刀原则“简单优于复杂”——说明应该尽可能地简单，只要不遗漏重要的信息。因此，应使我们建立的模型越简单越好。[1]即使知道其他变量可能会对Y有影响，我们也把这些次要的因素归入随机项u_i。

可能会由于上述一个或几个原因导致个人需求量偏离群体均值(即系统部分)。在随后的分析中，我们将会看到，随机误差项在回归分析中有着至关重要的作用。

5.5 样本回归函数

如何估计式(5-1)总体回归函数呢，也就是说，如何求得参数B_1,B_2呢？如果告知表5-1中的数据，即整个总体，那么这会是一项较为易做的工作。我们只要求出相对每一个X的Y的条件均值，然后再联立求解即可。不幸的是，在实际中，我们很少能拥有整个总体数据。通常，我们仅有来自总体的一个样本(回顾第1章，第2章有关总体和样本的讨论)。这里，我们的任务就是根据样本提供的信息来估计总体回归函数。如何完成这一工作呢？

表5-2 来自表5-1总体的随机样本

Y	X
49	1
45	2
44	3
39	4
38	5
37	6
34	7
33	8
30	9
29	10

表5-3 来自表5-1总体的另一随机样本

Y	X
51	1
47	2
46	3
42	4
40	5
37	6
36	7
35	8
32	9
30	10

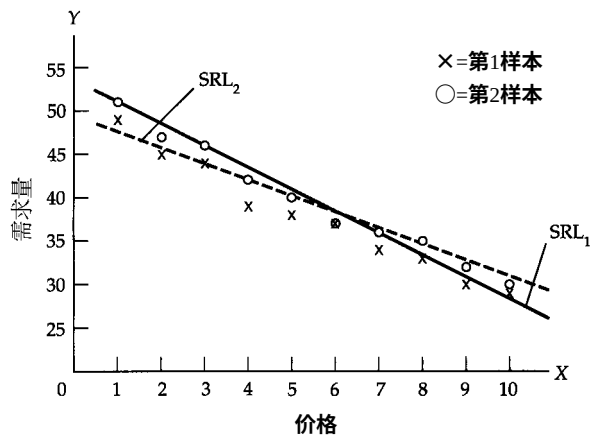


图5-3 样本回归直线

假设你重未见到过表5-1，你仅有表5-2提供的数据，这些数据是从表5-1中对每一个X随机

1 模型是现实的简单化。如果我们真想把现实转化为模型，那么模型将会变的非常复杂以至于没有任何实际的意义了。因此，在建立模型时，对现实进行提炼是很必要的。

抽取一个Y值得到的，与表5-1不同的是，此时对于每一个X值仅有一个Y值与之对应，现在面临的一个重要问题是：根据表5-2提供的样本数据，我们能够估计出相应于每一个X值，总体Y的均值吗？换句话说，我们能根据样本数据来估计总体回归函数吗？读者可能会想，我们或许不能“准确地”估计总体回归函数，因为存在抽样波动或是抽样误差(参见第4章)。为了更清楚地看此问题，我们假设有另一个来自表5-1总体的随机样本，见表5-3。对表5-2，5-3中的数据作图，得图5-3所示的散点图。

通过散点，可以清晰地得到两条很好地“拟合”了样本数据的直线，称之为样本回归线(sample regression lines, SRL)，这两条样本回归线哪一条代表了“真实的”总体回归直线呢？如果不看图5-1，那么我们将无法确定图5-3中哪一条直线代表了真实的总体回归线。如果再有一个样本，还可得到第三条样本回归线。恐怕每一条样本回归线都代表了总体回归线，但由于抽样的不同，每一条直线也最多是对真实总体回归线的近似。总之，我们可能从K个不同的样本中得到K条不同的样本回归直线，所有的这些样本回归线不可能都相同。

与从总体回归线得到总体回归函数类似，可用样本回归函数(sample regression function, SRF)来表示样本回归线。

$$\hat{Y}_i = b_1 + b_2 X_i \tag{5-3}$$

其中，

$\hat{Y}_i$ 表示总体条件均值，$E(Y|X_i)$的估计量；b_1表示B_1的估计量；b_2表示B_2的估计量；^ 读作“帽”。

前面讲过，估计量或样本统计量是用以表示如何估计总体参数的公式。估计量的某一取值称为估计值(回顾第4章中有关点估计量与区间估计量的讨论)。

从图5-3的散点图中不难看出：并非所有的样本数据都准确地落在各自的样本回归线上。因此，与建立随机总体回归函数一样，我们需要建立随机的样本回归函数：

$$Y_i = b_1 + b_2 X_i + e_i \tag{5-4}$$

其中，e_i是u_i的估计量。

我们称e_i为残差项(residual term)，或简称为残差(residual)。从概念上讲，它与u_i类似，可看做u_i的估计量。样本回归函数中生成e_i的原因与总体回归函数中生成u_i的原因相同。

总之，回归分析的主要目的是根据样本回归函数

$$Y_i = b_1 + b_2 X_i + e_i$$

来估计总体回归函数，

$$Y_i = B_1 + B_2 X_i + u_i$$

因为，通常我们的分析是根据来自某一总体的单独的一个样本。但是，由于抽样的不同，所以对总体回归函数的估计仅仅是近似估计。见图5-4。

对于图中的某个给定X_i，有一个(样本)观察值Y_i与之对应。利用样本回归函数形式，观察值Y_i可以表示为：

$$Y_i = \hat{Y}_i + e_i \tag{5-5}$$

利用总体回归函数形式，观察值Y_i可以表示为：

$$Y_i = E(Y \mid X_i) + u_i \tag{5-6}$$

显然，在图5-4中，$\hat{Y}_i$过高地估计了真实均值$E(Y \mid X_i)$。同样的道理，对图中点A右侧任一Y值，样本回归函数低估了真实的总体回归函数。但是，读者很容易发现：由于存在抽样波动，这些过高或过低的估计是不可避免的。

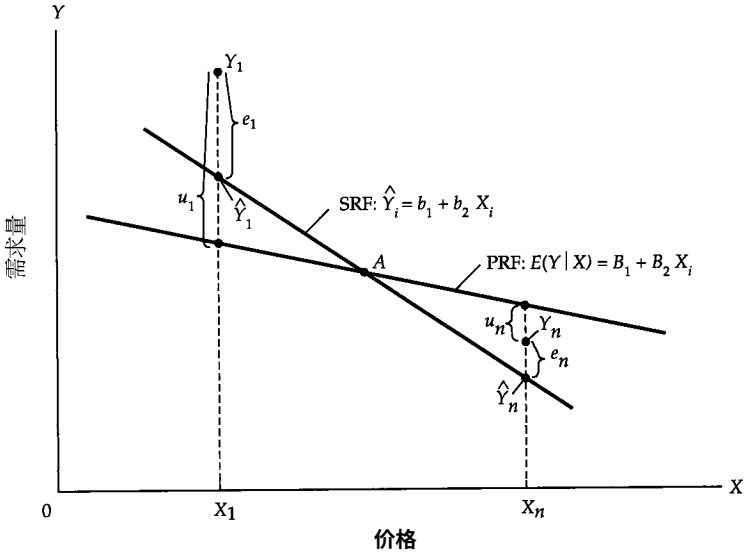


图5-4 总体回归线与样本回归线

一个重要的问题是：既然样本回归函数仅仅是总体回归函数的近似，我们能否找到一种方法(或过程)能够使这种近似尽可能接近真实值？也就是说，一般情况我们很难获得整个总体的数据，那么如何建立样本回归函数，使得 b_1,b_2尽可能接近B_1,B_2呢？在5.8节中将会看到，确实可以找到一个“最适合”的样本回归函数，它将尽可能“忠实地”反映总体回归函数。的确，我们总是幻想可以这样做，即使实际上无法确定真实的总体回归函数。

5.6 “线性”回归的特殊含义

本书主要关注诸如式(5-1)所描述的线性模型。由于对“线性”这一概念有两种不同的解释，因此，弄清楚线性的确切含义是非常必要的。

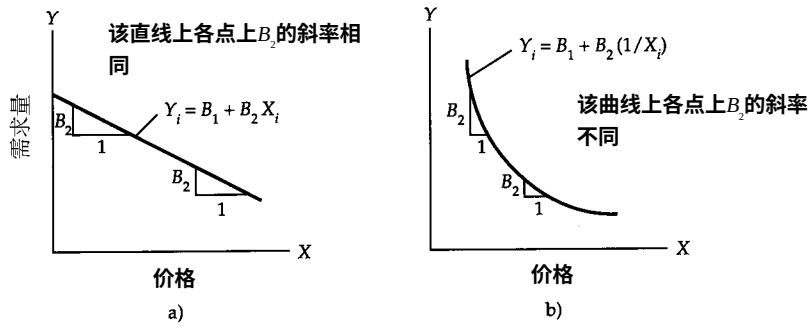


图 5-5

a)线性需求曲线 b)非线性需求曲线

5.6.1 解释变量线性

线性的第一个，也是最“本质”的含义是应变量的条件均值是自变量的线性函数。例如式(5-1)、式(5-2)以及相应的样本形式式(5-3)和式(5-4)。[1]按照这种解释，下面的函数形式不是线性的：

1 函数$Y=f(X)$称为线性的，如果(1)X仅以一次方的形式出现；也就是说，诸如 X^2, $\sqrt{X}$ 等都排除在外。(2)X不与其他变量相乘或相除(例如， XY, X/Y,其中Y是另一变量)。

$$E(Y)=B_1+B_2X_i^2 \tag{5-7}$$

$$E(Y)=B_1+B_2\frac{1}{X_i} \tag{5-8}$$

因为，在式(5-7)中X以平方的形式出现，而在式(5-8)中X以其倒数形式出现。由于回归模型中的解释变量线性，所以解释变量每变动一单位，被解释变量的变化率为一常量，也就是说，斜率保持不变。但对于非线性回归模型，斜率是变化的。[1]这一点从图5-5中可清楚看到。

图5-5表明，对于式(5-1)回归方程，无论X的值如何变化，斜率，即$E(Y)$的变化率保持不变，为B_2。但是，对于式(5-8)回归方程，Y的均值的变化率随回归线上的点的不同而变化，实际上，这里的回归线是一条曲线。[2]

5.6.2 参数线性

线性的第二解释是因变量的条件均值是参数B_S的线性函数；在这个线性函数中，解释变量之间并不一定是线性的。与解释变量线性函数类似，函数称为参数线性的，如果参数，比如说 B_2，仅以1次方形式出现。按照这一定义，模型(5-7)(5-8)均为线性模型。因为B_1,B_2以线性形式进入模型，而与变量是否呈线性无关。但是，下面的模型就是非线性的，因为B_2以平方的形式出现：

$$E(Y)=B_1+B_2^2X_i \tag{5-9}$$

在本书中，我们主要关注参数线性的模型。因此，从现在起，线性回归(linear regresion)是指参数线性的回归(即参数仅以一次方的形式出现在模型中)，而解释变量并不一定是线性的。[3]

5.7 从双变量回归到多元线性回归

到目前为止，我们仅考虑了双变量回归模型(two-variable regression)，或者简单地说是回归模型。在这类模型中，应变量仅是一个解释变量的函数。我们仅仅是通过双变量模型来介绍回归分析的基本思想，很容易将回归的概念推广到应变量是多个解释变量函数的情况。举个例子，若对widget的需求是价格(X_2)、消费者收入(X_3)以及竞争产品价格(X_4)的函数的话，那么，我们写出推广的需求函数：

$$E(Y)=B_1+B_2X_{2i}+B_3X_{3i}+B_4X_{4i} \tag{5-10}$$

即：$E(Y)=E(Y\mid X_{2i},X_{3i},X_{4i})$

式(5-10)就是多元线性回归(multiple linear regression)的一个例子。这个回归方程中有不止一个的自变量或解释变量用以解释应变量的行为。该模型表明：对 widget需求量的条件均值是价格、消费者收入及竞争产品价格的线性函数。单个消费者的需求函数(即随机的总体回归函数)为：

$$\begin{aligned}Y_i&=B_1+B_2X_{2i}+B_3X_{3i}+B_4X_{4i}+u_i\\&=E(Y)+u_i\end{aligned} \tag{5-11}$$

式(5-11)表明：由于随机误差项u_i的存在，个人需求量不同于群体的平均需求量。正如前面讲过的，即使在多元回归分析中，也需引进误差项，因为我们不能把所有可能影响需求量的因素都考虑进去。

---

1 考虑函数$Y=a+bX$，$\mathrm{d}Y/\mathrm{d}X=b$。再考虑函数$Y=a+bX^2$，$\mathrm{d}Y/\mathrm{d}X=2bX$，它不是常数，其值随$X$的变化而变化。$\mathrm{d}Y/\mathrm{d}X$($Y$对$X$的导数)度量了每单位$X$的变化所引起的$Y$的变化率。几何上，导数是函数的斜率。

2 学过微积分的读者知道在线性模型中，斜率(即 Y对X的导数)为一常数，等于 B_2，但在非线性模型中，式(5-8)，斜率等于$-B_2(1/X_i^2)$，显然，该斜率值依赖于X值的变化，因而它不是一个常数。

3 这也并不意味着像式(5-9)那样的非线性模型不能被估计，或是在实际中不能运用。事实上，在高级经济计量学教程中，对这类模型有很深地研究。

注意：式(5-10)和式(5-11)都是参数线性的，因此，它们都是线性回归模型。而进入模型的解释变量本身不需要是线性的。不过在 widget一例中，却是线性的。

5.8 参数的估计：普通最小二乘法

前面已经讲过，必须根据样本回归函数来估计总体回归函数，因为，在实际中，我们仅有来自某一总体的一两个样本。那么，如何估计总体回归方程呢？又如何判定估计的总体回归方程是否是真实的总体回归方程的一个“好的”估计呢？我们将在本章回答第一个问题，至于第二个问题，暂时放一放，留在下一章解答。

为了介绍总体回归方程估计的基本思想，我们考虑最简单的线性回归模型，即双变量线性回归模型(即研究因变量 Y与一个自变量 X之间的关系)。在第 7章，我们将把分析扩展到多元回归模型，在那里，将研究应变量与多个自变量之间的关系。

5.8.1 普通最小二乘法

虽然，有若干不同的方法来求样本回归函数(即真实总体回归函数的估计量)，但是，在回归分析中，用得最为广泛的方法是最小二乘法(ordinary least squares, OLS)，[1]一般称为普通最小二乘法(OLS)。我们将交错使用术语“最小二乘法”与“普通最小二乘法”。在介绍这种方法之前，我们先解释一下最小二乘原理(least squares principle)。

最小二乘原理

回顾式(5-2)所描述的双变量总体回归方程：

$$Y_i=B_1+B_2X_i+u_1$$

由于总体回归方程不能直接观察(为什么？)，我们用下面的样本回归函数来估计它。

$$Y_i=b_1+b_2X_i+e_i$$

因而，

$$\begin{aligned} e_i &= \text{实际的}Y_i - \text{估计的}Y_i \\ &= Y_i - \hat{Y}_i \\ &= Y_i - b_1 - b_2X_i \qquad [\text{利用式(5-3)}] \end{aligned} \tag{5-12}$$

从式(5-12)可以看出：残差是Y_i的真实值与估计值之差，而后者可从式(5-3)中得到。这一点可从图5-4中清晰地看到。

估计总体回归函数的最优方法是，选择 B_1，B_2的估计量b_1，b_2，使得残差e_i尽可能的小。虽然，可用几种不同的方法完成上述过程，但在回归分析中，用的最为广泛的方法之一是普通最小二乘法(ordinary least squares, OLS)，即选择参数 b_1，b_2，使得全部观察值的残差平方和(residual sum of squares, RSS)最小。

用数学形式表示为：

$$\begin{aligned} \min: \sum e_i^2 &= \sum (Y_i - \hat{Y}_i)^2 \\ &= \sum (Y_i - b_1 - b_2X_i)^2 \end{aligned} \tag{5-13}$$

总之，最小二乘原理就是所选样本回归函数使得所有 Y的估计值与真实值差的平方和最

1 除了名字以外，这种方法并没有什么普通的。事实上，我们将会看到，这种方法有一些特殊的性质。之所以称为普通最小二乘法是因为还有另一种方法，称之为广义最小二乘法(GLS)，普通最小二乘法是它的一个特例。

小。[1]

在阐述如何实现这个最小化过程之前，先来看一个具体的例子。

5.8.2 实例分析：对widget的需求

让我们再来看widget一例中的的需求表(表5-2)，为了方便起见，我们将该表加以复制并加以其他的计算，得到表5-4。回忆一下，表5-4给出了来自(总体)需求表5-1的随机样本。根据表5-4中的数据作图5-6。该图表明：实际的样本回归直线以最小二乘法的方式拟合了样本数据(这种方法的内在机理我们稍后会加以解释)。真实 Y的值与估计的Y值的垂直距离即为 e_i。在OLS估计中，这些垂直距离平方和即为 RSS($\sum e_i^2$)，见式(5-13)。普通最小二乘法就是使 RSS最小化的过程。

从式(5-13)可以看出，RSS($\sum e_i^2$)是b_1和b_2的函数。给定一组数据，比如表5-4提供的数据，选择不同的b_1，b_2值将会得到不同的e_S，因而使得RSS不同：只需随意地旋转图5-6中的样本回归函数，就可以看出。每一次旋转，都将得到一个不同的截矩(b_1)和一个不同的斜率(b_2)。(试着画出经过散点图的几条回归线)

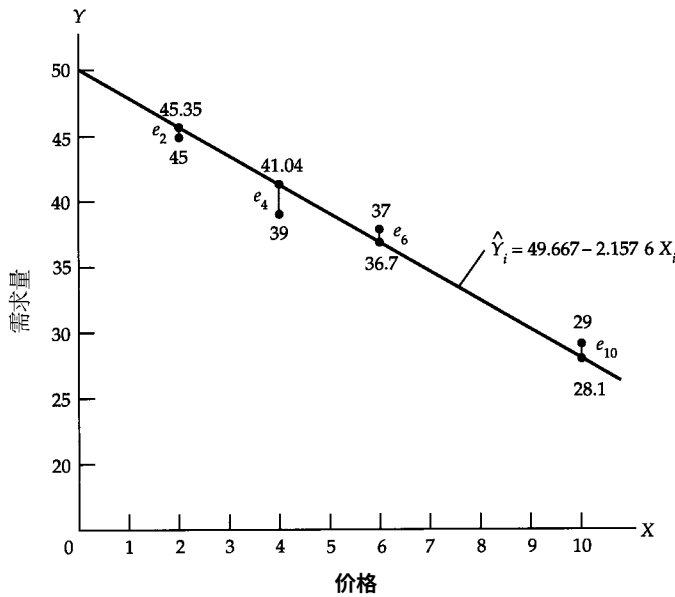


图5-6 真实Y值与估计Y值之差。普通最小二乘法(OLS)是使离差的平方和最小

这是否意味着给定一组数据，我们无法拟合一个确定的样本回归函数呢？也就是说，不可能求得惟一的b_1，b_2使得RSS最小？其实这正是OLS法的魅力所在。因为，对于一组数据，它给出了惟一的b_1，b_2。根据表5-4提供的数据，求的b_1，b_2的值分别为49.667 0和－2.157 6，见图5-6。(回顾：估计量的某一取值称为估计值)。

在实际中，如何确定这些值呢？这仅仅是一个的数学问题，需要用到差分知识。这里不再详细讨论，可以证明，b_1，b_2是使式(5-13)中的RSS最小化，求解下面的两个联立方程得到的：[2]

1 注意：e_i越小，其平方和也就越小。这里考虑 e_i的平方而非e_i，是因为这样一个过程可以避免残差的符号问题。需要提醒一点的是e_i的符号可正可负。

2 本章最后的附录中给出了详细的证明过程。

表5-4 widgets一例的最小二乘计算

	Y_i	X_i	X_iY_i	X_i^2	x_i	y_i	y_i^2	x_i^2	x_iy_i	$\hat{Y}_i$	e_i	e_i^2	e_iX_i
	49	1	49	1	−4.5	11.2	125.44	20.25	−50.4	47.509	1.4909	2.2228	1.4909
	45	2	90	4	−3.5	7.2	51.84	12.25	−25.2	45.352	−0.3519	0.1238	−0.7038
	44	3											
	39	4											
	38	5											
	37	6											
	34	7											
	33	8											
	30	9	270	81	5.5	−7.8	60.84	12.25	−27.3	30.248	−0.2484	0.0617	−2.2356
	29	10	290	100	4.5	−8.8	77.44	20.25	−39.6	28.091	0.9091	0.8265	9.0910
总和	378	55	1901.0	385	0	0	960.60	82.50	−178.0	378.00*	0*	9.9515	0*

斜率: $\hat{b}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{-178}{82.5} = -2.1576$

截距: $\hat{b}_1 = \bar{Y} - \hat{b}_2\bar{X} = 37.8 - (-2.1576)(5.5) = 49.667$

注：*表示四舍五入后的近似值

$\bar{Y} = \frac{378}{10} = 37.8$

$\bar{X} = \frac{55}{10} = 5.5$

$x_i = (X_i - \bar{X})$ $y_i = (Y_i - \bar{Y})$

$\hat{Y}_i$ 根据式(5-3)计算

表中空缺部分由读者完成

$$\sum Y_i = mb_1 + b_2 \sum X_i \tag{5-14}$$

$$\sum Y_i X_i = b_1 \sum X_i + b_2 \sum X_i^2 \tag{5-15}$$

其中，n为样本容量，这些联立方程称之为正规方程(normal equation)。

在式(5-14)和(5-15)中，参数b_s是未知的，但可以根据样本求得变量Y和X的平方和，交叉乘积和等等。现求解这些联立方程(运用中学学过的代数运算)，求得 b_1，b_2：

$$b_1 = \bar{Y} - b_2 \bar{X} \tag{5-16}$$

它是总体截距B_1的估计量。

$$\begin{aligned} b_2 &= \frac{\sum x_i y_i}{\sum x_i^2} \\ &= \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \\ &= \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2} \end{aligned} \tag{5-17}$$

它是总体斜率B_2的估计量。

式中：$x_i = X_i - \bar{X}, y_i = Y_i - \bar{Y}$ ；即小写字母代表了变量与其均值的偏差。

式(5-17)给出了计算斜率的几种不同方法，你可以选择适当的方法求解。不过，在计算机时代，任何一种方法在计算时都不成问题。

式(5-16)、(5-17)给出的估计量称为最小二乘估计量(OLS estimators)，因为它们是用OLS法求得的。

回到widget一例中，表5-4提供了计算OLS估计量b_1，b_2所需的全部数据，我们得到估计的需求函数：

$$\hat{Y}_i = 49.667 - 2.157\,6 X_i \tag{5-18}$$

它就是图5-6描述的样本回归函数。

5.8.3 对估计的需求函数的解释

对估计出的需求函数解释如下：在其他条件保持不变的情况下，如果 widget的价格上升，比如说上升 1美元，平均而言，则对 widget的需求量下降 2.16个单位(也即需求曲线的斜率－2.16)。类似地，如果widget的价格为0，则平均的需求量为49.7个单位。(也即需求曲线的截距为49.7个单位)。当然，这只是截距这一概念在数学上机械地解释。在现实中，产品的价格绝对不为0，除非它是“自由享用”商品，比如空气(但是注意：没有污染的空气却不是免费的)。通览全书，你会发现截距一般都没有什么特殊的经济含义。

在继续下面内容之前，我们先给出普通最小二乘估计量的一些有趣的性质。

(1) 运用OLS法得出的样本回归线经过样本均值点，即

$$\bar{Y}_i = b_1 + b_2 \bar{X} \tag{5-19}$$

(2) $\bar{e} = \sum e_i / n$ ，即残差的均值总为零，我们可以利用这条性质检验计算的准确与否。(证明见习题5.11)。

(3) 对残差与解释变量的积求和，其值为零；也就是说，这两个变量不相关(参见第 2章中相关的定义)。

$$\sum e_i X_i = 0 \tag{5-20}$$

这条性质也可用来检查最小二乘法计算的结果。

5.9 实例

既然已经讨论了OLS法，并且了解了如何对总体回归函数进行估计，现在我们来看一些具体实例。

例5.1

有一个由212个家庭组成的随机样本，根据OLS法得到的回归方程如下：[1]

$$\widehat{Y}_i = -1.924 + 0.19X_i$$

其中，Y表示家庭支付的联邦收入税，X表示家庭收入，它们均以美元度量。

在这个例子中，b_1=－1.924,b_2=0.19。回归结果表明：在其他条件不变的情况下，家庭收入每增加1 000美元，平均而言，税收将增加190美元。也可以说，家庭收入每增加1个单位，比如说，1美元，税收将平均地增加0.19美元或19美分。只不过在此例中，收入是以数以千记的美元来度量的。

前面已经提到过，在大多数情况下，截距没有什么明显的经济含义，本例亦如此。字面上解释截距就是家庭收入为零时的税赋。这并没有什么经济意义，除非我们把这个数字解释为“负的”收入税；换句话说，就是在此情况下，政府实际付给家庭1 924美元！

表5-5 S.A.T分数

美国高年级学生平均学生智能测试结果，1967～1990，1967～1971的数据为估计值

	词汇			数学		
年度	男生	女生	总计	男生	女生	总计
1967	463	468	466	514	467	492
1968	464	466	466	512	470	492
1969	459	466	463	513	470	493
1970	459	461	460	509	465	488
1971	454	457	455	507	466	488
1972	454	452	453	505	461	484
1973	446	443	445	502	460	481
1974	447	442	444	501	459	480
1975	437	431	434	495	449	472
1976	433	430	431	497	446	472
1977	431	427	429	497	445	470
1978	433	425	429	494	444	468
1979	431	423	427	493	443	467
1980	428	420	424	491	443	466
1981	430	418	424	492	443	466
1982	431	421	426	493	443	467
1983	430	420	425	493	445	468
1984	433	420	426	495	449	471
1985	437	425	431	499	452	475
1986	437	426	431	501	451	475
1987	435	425	430	500	453	476
1988	435	422	428	498	453	476
1989	434	421	427	500	454	476
1990	429	419	424	499	455	475

资料来源：The College Board .The *New York Times,* Aug .28,1990,p.B-5.

1 参见Richard G.Lipsey,Peter O.Steiner,Douglas D.Purvis，*Economics*, 8 th ed., Harper & Row, New York,1987。其中的符号有所变化。注意：为了简便，有时用“OLS回归”表示通过OLS法得到的回归函数。

例5.2

表5-5给出1967～1990年期间男女学生的词汇及数学分数。假设我们想根据男生的数学分数(X)来预测男生的词汇分数(Y)。利用表5-5提供的数据，运用OLS法，得到下面的回归结果：

$$\hat{Y}_t = -380.48 + 1.641\ 8X_t \tag{5-22}$$

因为该数据为时间序列，所以把时间 t作为变量的下标。这里共有 24年的数据($t=24$)。

这个回归结果有什么意义呢？这里，斜率为 1.641 8，意指如果男生的数学分数每增加1分，平均而言，其词汇分数将增加 1.64分。这两个变量看似正相关。截距－380.48，没有什么实际意义，因为即使数学分数为零，词汇分数也不可能为负。当然，也有例外， Board大学在举行S.A.T考试中，如果考生数学分数为零，则作为惩罚，他或她的词汇分数将为－380分。

例5.3

女生的词汇分数(Y)与数学分数(X)之间的关系

根据表5-5提供的数据，运用OLS法，得到如下回归结果：

$$\hat{Y}_t = -336.06 + 1.698\ 5X_t \tag{5-23}$$

这里b_1=－336.06,b_2=1.698 5。即，当其他条件保持不变时，女生的数学分数每提高1分，平均而言，其词汇分数平将提高 1.70分，这个增加值稍大于男生的增加值。同样，截距－336也没有什么实质意义。

例5.4

奥肯定律

布鲁金斯学会主席，前总统经济顾问委员会主席奥肯 (Arthur Okun)根据美国1947-1960年的数据，得到如下回归方程，称之为奥肯定律：

$$\hat{Y}_t = -0.4(X_t - 2.5)$$

其中，Y_t表示失业率的变动率(百分数)，

X_t表示实际产出的增长率(百分数)，用实际GNP度量，

2.5是对美国历史地观察得到的长期产出增长率。

在这个回归方程中，截距为零，斜率为－0.4。奥肯定律是说实际GNP的增长每超过2.5%一个百分点，失业率将降低0.4个百分点。

奥肯定律被用来预测为使失业率减少到一定的百分点而所需的实际 GNP的增长率。因此，实际GNP的增长率为5%时，将使失业率减少一个百分点，或者说若使增长率达到7.5%，则需减少失业率2个百分点(验证这些结论)。

通过这个例子，我们可以了解回归结果是如何被运用到政策当中去的。

例5.5

名义汇率与相对价格间的关系

参考习题1.7中的表1-3给出的数据。令 Y代表德国马克对美元的汇率 (GM/$)，$X$代表美国与德国的消费者价格指数之比，即 X表示了两个国家的相对价格。利用表1-3给出的数据，得到如下结果：

$$\hat{Y}_t = 6.682 - 4.318 X_t \tag{5-25}$$

在这个回归方程中，斜率为－ 4.318，表明在1980～1994年期间，相对价格每上升一单位，平均而言，马克对美元的汇率下降 4.32个单位，也即美元将会贬值，因为一美元将兑换更少的德国马克。斜率的经济含义是：如果价格在美国上涨得比在德国上涨得快，则美国国内的消费者将转向消费德国产品，因此，对马克的需求上升；从而导致马克的增值。这就是购买力平价理论 (或称单一价格法则)的实质。

截距6.682表示若相对价格为零，则马克对美元汇率的均值为 6.682，即1美元可以兑换6.682德国马克。但是，这种解释也并没有什么经济意义。 (为什么？)

例5.6

MBA毕业生的基本年薪(ASP)与GMAT分数之间的关系(1994年)

参考习题5.17给出的数据，得出如下的回归结果：

$$\widehat{ASP}_i = -335\,983 + 648.08\text{GMAT}_i \tag{5-26}$$

从式(5-26)看， GMAT分数越高，则 MBA毕业生的基本年薪就越高。斜率648.08表示GMAT分数每提高1分，平均而言， MBA毕业生基本年薪将增加 648美元，负的截距没有实际意义，因为基本年薪不可能为负值

当然，除GMAT分数之外，还有其他一些影响 ASP的因素。在第7章中，我们将根据习题5.17的数据建立一个多元回归模型，从而考虑更多的变量。

关于计算

式(5-22)和(5-23)中的回归结果是根据表 5-5提供的数据，利用普通最小二乘法 [式(5-10)，(5-17)]计算得到的。显然手工计算将是繁锁而耗时的。如今，我们可以用统计软件很快地估计类似式(5-22)、(5-23)那样的回归模型。在第 1章中已经提到过，在本书中，我们会不断地使用这些统计软件。因此，在适当的地方，还将给出利用这些软件得出的回归结果，以便使读者更好地了解它们。

5.10 小结

本章介绍了回归分析的基本思想。从总体回归函数(PRF)开始，建立了线性总体回归函数的概念。这也是本书所关注的主要内容。线性回归是指参数线性的，而不论变量线性与否。我们还介绍了随机的总体回归函数，并且详细讨论了随机误差项的一些特性及作用。当然，总体回归函数只是建立在理论之上的，或者说是理想化的，因为，在实际中，我们仅仅可能从某一总体中获得一个或若干个样本。这也正是讨论样本回归函数(SRF)的重要意义之所在。

本章还讨论了如何获得样本回归函数。在此介绍了最常用的普通最小二乘法，并给出估计总体回归函数参数的相应计算公式。

我们用一个数值例子及若干具体实例阐述了普通最小二乘法。

下一个重要的工作是如何判定用OLS法得到的样本回归函数的优度，这也正是下一章所要讨论的内容。

习题

5.1 解释概念

(a) 总体回归函数(PRF) (b) 样本回归函数(SRF) (c) 随机的总体回归函数 (d) 线性回归模型 (e) 随机误差项(u_i) (f) 残差项(e_i) (g) 条件期望 (h) 非条件期望 (i) 回归系数或回归参数 (j) 回归系数的估计量

5.2 随机的总体回归函数与随机的样本回归函数有何区别？

5.3 评论“既然我们不能观察到总体回归函数，为什么还要研究它呢？”

5.4 判断正误并说明理由。

(a)随机误差项与u_i与残差项e_i是一回事。 (b)总体回归函数给出了对应于每一个自变量的因变量的值。 (c)线性回归模型意味着变量是线性的。 (d)在线性回归模型中，解释变量是原因，被解释变量是结果。 (e)随机变量的条件均值与非条件均值是一回事。 (f)式(5-22)中的回归系数B是随机变量，但式(5-4)中的回归系数b_S是参数。 (g)方程(5-1)中的斜率B_2度量了X每变动1单位，Y的倾斜度。 (h)在实际中，双变量回归模型没有什么用，因为应变量的行为不可能仅由一个解释变量来解释。

5.5 下表列出了若干对自变量与应变量。对每一对变量，你认为它们之间的关系如何？是正的？负的？还是无法确定？也就是说，其斜率是正还是负，或都不是？并说明理由。

应变量	自变量
(a) GNP	利率
(b) 个人储蓄	利率
(c) 小麦产出	降雨量
(d) 美国国防开支	前苏联国防开支
(e) 棒球明星本垒打的次数	其年薪
(f) 总统声誉	任职时间
(g) 学生第一年考试成绩平均分数	其S.A.T分数
(h) 学生经济计量学成绩	其统计学成绩
(i) 日本汽车的进口量	美国人均国民收入

5.6 判别下列模型是否为线性回归模型：

(a) $Y_i=B_1+B_2(1/X_i)$ (b) $Y_i=B_1+B_2\ln X_i+u_i$ (c) $\ln Y_i=B_1+B_2X_i+u_i$ (d) $\ln Y_i=B_1+B_2\ln X_i+u_i$ (e) $Y_i=B_1+B_2B_3X_i+u_i$ (f) $Y_i=B_1+B_2^3X_i+u_i$

注 自然对数表示以e为底的常用对数。

5.7 下表给出了每周家庭的消费支出Y(美元)与每周家庭的收入(美元)的数据。

每周收入(X)	每周消费支出(Y)
80	55,60,65,70,75
100	65,70,74,80,85,88
120	79,84,90,94,98
140	80,93,95,103,108,113,115
160	102,107,110,116,118,125
180	110,115,120,130,135,140
200	120,136,140,144,145
220	135,137,140,152,157,160,162
240	137,145,155,165,175,189
260	150,152,175,178,180,185,191

(a) 对每一收入水平，计算平均的消费支出，$E(Y|X_i)$，即条件期望值。 (b) 以收入为横轴，消费支出为纵轴作散点图。 (c) 在该散点图上，作出(a)中的条件均值点。 (d) 你认为X与Y之间，X与Y的均值之间的关系如何？ (e) 写出其总体回归函数及样本回归函数。 (f) 总体回归函数是线性的还是非线性的？

5.8 根据上题中给出的数据，对每一个X值，随机抽取一个Y值，结果如下：

Y	70	65	90	95	110	115	120	140	155	150
X	80	100	120	140	160	180	200	220	240	260

(a) 以Y为纵轴，X为横轴作图。 (b) 你认为Y与X之间是怎样的关系？ (c) 求样本回归函数？按照表5-4的形式写出计算步骤？ (d) 在同一个图中，作出样本回归函数以及从习题5.7中得到的总体回归函数。 (e) 总体回归函数与样本回归函数相同吗？为什么？

5.9 假定有如下的回归结果：

$$\hat{Y}_t = 2.6911 - 0.4795X_t$$

其中，Y表示美国的咖啡的消费量(每天每人消费的杯数)

X表示咖啡的零售价格(美元/磅)

t表示时间

(a) 这是一个时间序列回归还是横截面序列回归？ (b) 画出回归线。 (c) 如何解释截距的意义？它有经济含义吗？ (d) 如何解释斜率？ (e) 你能求出真实的总体回归函数吗？ (g) 需求的价格弹性定义为：价格每变动百分之一所引起的需求量变动的百分比，用数学形式表示为：

$$\text{弹性} = \text{斜率} \times \left(\frac{X}{\bar{Y}}\right)$$

即，弹性等于斜率与X与Y比值之积，其中X表示价格，Y表示需求量。根据上述回归结果，你能求出对咖啡需求的价格弹性吗？如果不能，计算此弹性还需要其他什么信息？

5.10 下表给出了消费者价格指数(1982～1984=100)及500支股票的标准普尔指数(基准指数：1941～1943 =10)

(a) 以CPI指数为横轴，S&P指数为纵轴作图。 (b) 你认为CPI指数与S&P指数之间关系如何？ (c)考虑下面的回归模型： $(S\&P)_t = B_1 + B_2 CPI_t + u_t$ 根据表中的数据，运用普通最小二乘法估计上述方程并解释你的结果。 (d)(c)中的结果有经济意义吗？ (e)你知道为什么1988年S&P指数下降了吗？

5.11 证明：$\sum e_i = 0$，从而证明：$\bar{e} = 0$

5.12 证明 $\sum e_i x_i = 0$

5.13 证明 $\sum e_i \hat{Y}_i = 0$；即残差e_i与Y_i的估计值之积的和为零。

5.14 下表给出了1988年9个工业国的名义利率(X)与通货膨胀率(Y)的数据，

国家	Y(%)	X(%)
澳大利亚	11.9	7.7
加拿大	9.4	4.0
法国	7.5	3.1
德国	4.0	1.6
意大利	11.3	4.8
墨西哥	66.3	51.0
瑞典	2.2	2.0
英国	10.3	6.8
美国	7.6	4.4

资料来源：Rudiger Dormbush，Stanley Fischer《宏观经济学》，第五版，1990，第652页。原始数据来自国际货币基金组织出版的《国际金融统计》。

(a)以利率为纵轴，通货膨胀率为横轴作图。 (b)用OLS法进行回归分析，写出求解步骤。(c)如果实际利率不变，则名义利率与通胀率的关系如何？即在 Y对X的回归中，斜率如何？对名义利率、通胀率及实际利率之间关系的讨论可参见有关的宏观经济学教材，比如 Dornbush与Fisher的《宏观经济学》，还可查阅以美国著名的经济学家 Irving Fisher命名的Fisher方程。

5.15 实际汇率(RE)定义为名义汇率(NE)与本国价格与外国价格之比的乘积。

因此，美国对德国的实际汇率为：

$$RE_{US}=NE_{US}(US_{CPI}/German_{CPI})$$

(a)利用习题1.7中表1-3给出的数据，计算RE_{US}。 (b)利用你熟悉的回归分析软件，对下面回归模型进行估计。

$$NE_{US}=B_1+B_2RE_{US}+u \tag{1}$$

(c)先验地，你预期名义汇率与真实汇率的关系如何？你可以从有关国际贸易和宏观经济学的书中查阅有关购买力平价的理论。 (d)回归的结果验证了你的先验预期吗？如果没有，可能的原因是什么呢？ *(e)估计如下形式的回归方程：

$$1nNE_{US}=A_1+A_21nRE_{US}+u \tag{2}$$

其中，1n表示自然对数，即以e为底的常用对数。(1)式的回归结果和(2)式的回归的结果相同吗？说明理由。

5.16 参考题5.10。下表给出了1990～1996年间的CPI指数与S&P500指数的数据。

(a)重复习题5.10(a)到(e)的各个问题。 (b)你认为估计的回归模型有什么不同？ (c)现将两组数据联合起来，再进行回归分析。 (d)三个回归模型存在显著差异吗？

年 份	CPI	指 数
1990	130.7	334.59
1991	136.2	376.18
1992	140.3	415.74
1993	144.5	451.41
1994	148.2	460.33
1995	152.4	541.64
1996	159.6	670.83

资料来源：总统经济报告，1997，CPI指数见表B-60，第380页；S&P指数见表B-93，第406页。

5.17 下表给出了美国30所知名学校的MBA学生1994年基本年薪(ASP)、GPA分数(从1到4共四个等级)、GMAT分数以及每年学费的数据。

(a)用双变量回归模型分析GPA是否对ASP有影响？ (b)用合适的回归模型分析GMAT分数是否与ASP有关系？ (c)每年的学费与ASP有关吗？你是如何知道的？如果两变量之间正相关，是否意味着进到最高费用的商业学校是有利的。 (d)你同意高学费的商业学校意味着高质量的MBA成绩吗？为什么？

表5-6 1994年MBA毕业生平均初职薪水

学 校	ASP/美元	GPA	GMAT	学费/美元
Harvard	102 630	3.4	650	23 894
Stanford	100 800	3.3	665	21 189
Columbian	100 480	3.3	640	21 400
Dartmouth	95 410	3.4	660	21 225
Wharton	89 930	3.4	650	21 050
Northwestern	84 640	3.3	640	20 634
Chicago	83 210	3.3	650	21 656
MIT	80 500	3.5	650	21 690

(续)

学 校	ASP/美元	GPA	GMAT	学费/美元
Virginia	74 280	3.2	643	17839
UCLA	74 010	3.5	640	14496
Berkeley	71 970	3.2	647	14361
Cornell	71 970	3.2	630	20400
NYU	70 660	3.2	630	20276
Duke	70 490	3.3	623	21910
Carriegie Mellon	59 890	3.2	635	20600
North Carolina	69 880	3.2	621	10132
Michigan	67 820	3.2	630	20960
Texas	61 890	3.3	625	8580
Indiana	58 520	3.2	615	14036
Purdue	54 720	3.2	581	9556
Case Western	57 200	3.1	591	17600
Georgetown	69 830	3.2	619	19584
Michigan State	41 820	3.2	590	16057
Penn State	49 120	3.2	580	11400
Southern Methodist	60 910	3.1	600	18034
Tulane	44 080	3.1	600	19550
Illinois	47 130	3.2	616	12628
Lowa	41 620	3.2	590	9361
Minnesota	48 250	3.2	600	12618
Washington	44 140	3.3	617	11436

附录5A 最小二乘估计量的推导

从式(5-13)开始：

$$\sum e_i^2 = \sum (Y_i - b_1 - b_2 X_1) \tag{5A-1}$$

利用偏微分，得

$$\partial \sum e_i^2 / \partial b_1 = 2 \sum (Y_i - b_1 - b_2 X_1) \quad (-1) \tag{5A-2}$$

$$\partial \sum e_i^2 / \partial b_2 = 2 \sum (Y_i - b_1 - b_2 X_i) \quad (-X_i) \tag{5A-3}$$

根据最优化的一阶条件，令上述两式为零，于是有：

$$\sum Y_i = nb_1 + b_2 \sum X_i \tag{5A-4}$$

$$\sum Y_i X_i = b_1 \sum X_i + b_2 \sum X_i^2 \tag{5A-5}$$

即为正文中的式(5-14)和式(5-15)。

联力求解这两个方程，即得式(5-16)和(5-17)。

第 6 章

双变量模型：假设检验

在前面一章里，我们介绍了最小二乘法。通过表 5-2给出的样本的数据，运用最小二乘法，求得下面的需求函数：

$$\hat{Y}_i = 49.667 - 2.157\,6X_i \tag{5-18}$$

这是统计推断的估计阶段，我们现在把注意力转向统计推断的另一个阶级 —— 假设检验。在此，提出一个重要问题：式(5-18)给出的估计回归直线的“优度”如何？也就是说，怎样判别它确实是真实的总体回归函数的一个好的估计量呢？如何仅仅根据表 5-2给出的一个样本，来确定估计的回归函数(即样本回归函数)确实是真实总体回归函数的一个好的近似呢？

我们无法清楚地回答这个问题，除非对总体回归函数知道的更多一些。式(5-2)表明，Y_i依赖于X_i与u_i。现在假设X_i值是给定的或是已知的 —— 回顾在第5章中，我们的分析是条件回归分析，是以给定X_s为条件。简言之，我们把X看作是非随机的。随机误差项u当然是随机的(为什么？)。由于Y的生成是在随机误差项(u)上加上一个非随机项(X)，因而Y也就变成了随机变量。所有这些意味着：只有假定随机误差项是如何生成的，才能判定样本回归函数对真实回归函数拟合的好坏。到目前为止，在求普通最小二乘估计量的过程中，没有涉及 u_i是如何生成的，因为OLS估计量的生成与随机误差项的假定无关。但是，在对样本回归函数进行假设检验时，如果不知道误差项的生成过程，假设检验将无法进行，随后将会看到，只有对 u_i的生成做一些特殊的假定，才能完成假设检验。这正是将要讨论的古典线性回归模型(Classical Linear Regression Model, CLRM)。我们将沿用第5章中介绍的双变量回归模型来解释其基本思想。在第7章中，我们还将把这一思想推广到多元回归模型之中。

6.1 古典线性回归模型

古典线性回归模型有如下一些基本假定：

A6.1 解释变量(X)与扰动误差项不相关。但是，如果X是非随机的，(即其值为固定数值)，则该假定自动满足。

这个假定对我们来说并不陌生，在第 5章中，我们已经明确，回归分析是条件回归分析，是以给定X值为条件的。从根本上说，我们一直假定 X是非随机的。这个假定主要是为了处理第15章中的联立方程回归模型的问题。

A6.2 扰动项的期望或均值为零。即，

$$E(u_i)=0 \tag{6-1}$$

该假定表明：平均地看，随机扰动项对 Y_i 没有任何影响，也就是说，正值与负值相互抵消。图6-1给出了它的直观图形。

A6.3 同方差(homoscedastic)假定，即每个 u_i 的方差为一常数 σ^2。

$$\mathrm{Var}(u_i)=\sigma^2 \tag{6-2}$$

图6-2a给出了这一假定的几何图形。该假定可简单地理解为,与给定 X 相对应的每个 Y 的条件分布同方差；也即,每个 Y 值以相同的方差,分布在其均值周围。[1]如果不是这种情况,则称为异方差,见图6-2b。b图表明,每个 Y 的方差不同。(与图6-2a对照)。古典线性回归模型对方差的假定见图6-2a。

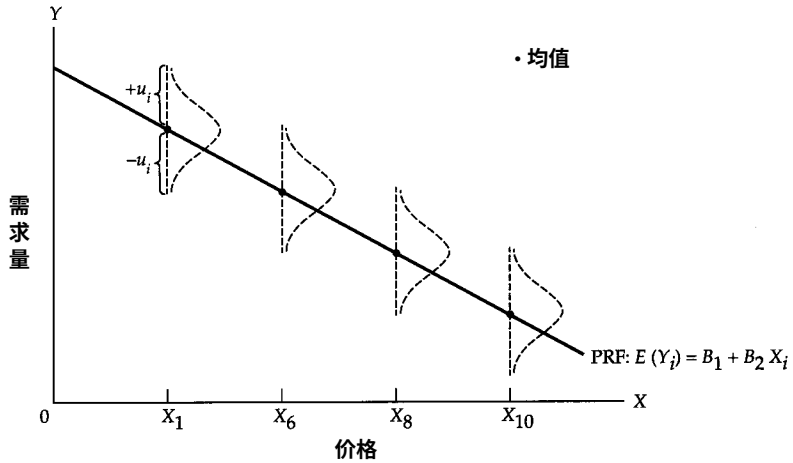


图6-1 扰动项的条件分布

A6.4 无自相关(no autocorrelation)假定，即两个误差项之间不相关。其数学数形式为,

$$\mathrm{cov}(u_i,u_j)=0 \quad i\neq j \tag{6-3}$$

这里，cov表示协方差(见第2章)，i 和 j 表示任意的两个误差项。(注：如果 $i=j$，则式(6-3)就给出了的方差的表达式)。

图6-3给出了自相关的几种图形。假定6.4表明：两误差项之间没有系统的关系。如果某一个误差项 u 大于其均值，并不意味着另一个误差项也在均值之上(对正相关而言)，或者，如果一个误差项小于均值，也并不意味着另一个误差项一定在均值之上，反之亦反(对于负相关而言)。简言之，无自相关假定表明误差项 u_i 是随机的。

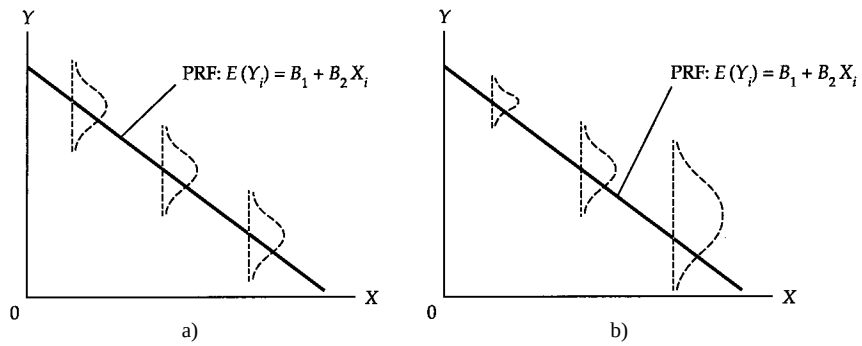


图 6-2
a)同方差 b)异方差

1 由于 X 值是假设给定的或是非随机的，因此，Y 中惟一变化的部分来自于 u。因此，给定 X_i，u_i 与 Y_i 同方差。简言之，u_i 的(条件)方差等于 Y_i 的(条件)方差 σ^2。但是，注意：在第2章中讨论过，Y 的非条件方差为 $E[Y_i-E(Y)]^2$。与之相对，Y 的样本非条件方差为 $\sum(Y_i-\bar{Y})/(n-1)$ 。

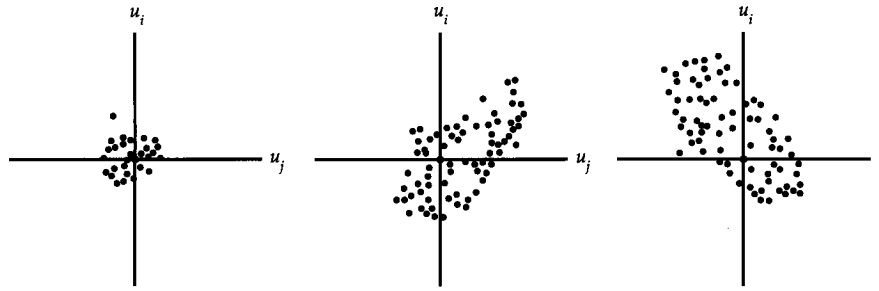


图6-3　自相关

a)无自相关　b) 正的自相关　c)负的自相关

读者或许对这些假定感到迷惑，为什么需要这些假定呢？它们的现实意义如何呢？如果这些假定不为真，情况又会怎样呢？如何知道某一回归模型确实满足所有这些假定呢？这些问题显然是很重要的，但是，在这一专题的开始阶段，我们不能对所有这些问题都给出满意的答案。事实上，本书的整个第三部分都是围绕着(古典线性回归模型的)一个或若干个假定不满足时，会发生什么情况而展开的。需要记住的是：对任何一门科学的探求，我们都会做一些假定，因为这样会有助于我们逐步地建立专题问题，而不是因为这些假定是现实所必需的。这里，类比法将会对我们的学习有所帮助。经济系的学生在学习不完全竞争模型之前，总是先学完全竞争的模型。因为对完全竞争模型的学习可以使学生更好地理解不完全竞争模型，而并不是因为完全竞争模型是现实所必需的。此外，也有一些市场是完全竞争的，例如股票市场和外汇市场。

6.2　普通最小二乘估计量的方差与标准差

上述这些假定的一个直接结果是：它们使我们能够估计式(5-16)和(5-17)给出的这些OLS估计量的方差及标准差。在第4章已经讨论过基本的估计理论，包括(点)估计量，抽样分布以及估计量的方差及标准差的概念。回顾这些知识，我们知道式(5-10)和(5-11)式给出的OLS估计量是随机变量，因为它们的值随样本的不同而变化。很自然地，我们想了解这些估计量的抽样差异性，即它们随着样本的变化是如何变化的。这种抽样差异性通常由估计量的方差或其标准差来度量。式(5-16)，式(5-17)给出的OLS估计量的方差(variance)及标准差(standard error)：[1]

$$\text{var}(b_1) = \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2 \tag{6-4}$$

(注：这个公式中既有小写的x，又有大写的X。)

$$se(b_1) = \sqrt{\text{var}(b_1)} \tag{6-5}$$

$$\text{var}(b_2) = \frac{\sigma^2}{\sum x_i^2} \tag{6-6}$$

$$\text{se}(b_2) = \sqrt{\text{var}(b_2)} \tag{6-7}$$

其中，var表示方差，se表示标准差，σ^2是扰动项u_i的方差。根据同方差假定，每一个 u_i具有相同的方差σ^2。

一旦知道了σ^2，就很容易计算等式右边的项，从而可以求得 OLS估计量的方差与标准差。同方差σ^2，由下式来估计：

1　证明见Damodar N, Gjarati, *Basic Econometrics*，3d ed., McGraw-Hill, New York，1995，pp.95-96。

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2} \tag{6-8}$$

其中，$\hat{\sigma}^2$ 是σ^2的估计量(回顾一下，用符号^表示一个估计量)，$\sum e_i^2$ 是残差平方和(RSS)，即Y的真实值与估计值差的平方和，$\sum (Y_i - \bar{Y})^2$ 。

$(n-2)$称为自由度，在第3章中已经讲过，它可以简单地看作是独立观察值的个数。[1]

一旦计算出e_i，就很容易求得 $\sum e_i^2$ ，同时有，

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} \tag{6-9}$$

即，$\hat{\sigma}^2$ 的正平方根称为估计值的标准差(standard error of the estimate)或是回归标准差(standard error of the regression)，它是Y值偏离估计的回归直线的标准方差。[2]估计值的标准差通常用作对估计回归线的拟合优度(goodness of fit)的简单度量，我们将在6.6节中详细讨论。

6.2.1 Widget一例中的方差和标准差

利用上述公式，计算widget一例的方差及标准差，见表6-1，必要的原始数据已在表5-4中给出。

6.2.2 widget需求函数小结

估计的对widget的需求函数如下：

$$\hat{Y}_i = 49.667\,0 - 2.157\,6X_i \tag{6-16}$$
$$\text{se}=(0.746\,4)\quad(0.120\,3)$$

其中，括号内的数字表示估计的标准差。有些时侯，回归结果是以上面的形式给出(在 6.7节中，更多地使用了这种形式)。

表6-1 widget一例中计算

估计量	公式	结果	式
$\hat{\sigma}^2$	$\Sigma \frac{e_i^2}{n-2} = \frac{9.551\,5}{8}$	1.193 9	(6.10)
$\hat{\sigma}$	$\sqrt{\hat{\sigma}^2} = \sqrt{1.193\,9}$	1.092 6	(6.11)
var (b_1)	$\frac{\Sigma X^2 \cdot \sigma^2}{n\Sigma x_i^2} = \frac{(385)(1.193\,9)}{10(82.5)}$	0.557 2	(6.12)
se (b_1)	$\sqrt{\text{var}(b_1)} = \sqrt{0.557\,2}$	0.746 4	(6.13)
var (b_2)	$\frac{\sigma^2}{\Sigma x_i^2} = \frac{1.193\,9}{82.5}$	0.014 5	(6.14)
se (b_2)	$\sqrt{\text{var}(b_2)} = \sqrt{0.014\,5}$	0.120 3	(6.15)

注：原始数据见表5-4，在计算估计量的方差时，用σ^2的无偏估计量 $\hat{\sigma}^2$ 代替他。

该表达式直接给出了估计的参数值及其标准差。例如，它告诉我们需求函数的斜率(即，价格变量的系数)为－2.157 6，其标准差为0.120 3——它度量了来自于不同样本的b_2差异性。如何利用这一结果呢？比如，我们能说计算 b_2位于真实值B_2的某个误差范围内吗？如果可以，那么，

1 注意，只有计算出 $\hat{Y}_i$ ，才能求出e_i，但是要计算e_i，必须先求b_1和b_2。在估计这两个未知参数时，我们失去2个自由度，因此，虽然有几个观察值，但自由度仅为 $(n-2)$。

2 注意回归标准差($\hat{\sigma}$)与样本标准差(S_y)的区别。后者是依据Y的均值来度量的，计算公式为：

$$S_y = \sqrt{\sum \frac{(Y_i - \bar{Y})^2}{n-1}}$$

而前者是依据由样本回归得到的估计值($\hat{Y}_i$)来度量的，参见脚注1。

在某一置信区间内，能否确定样本回归函数(6-16)对真实总体回归函数的拟合优度？当然，这是假设检验的内容。

但是，在讨论假设检验之前，需要了解更多的理论。特别地，既然 b_1和b_2是随机变量，必须求得它们的抽样(Sampling)或概率分布(Probability distribution)：回顾第3章、第4章讲到的随机变量的概率密度。在6.5节中，我们将会看到，一旦确定了这两个估计量的抽样分布，那么假设检验就是举手之劳的事了。这里，我们先要回答的一个重要问题：为什么要用 OLS法？

6.3 普通最小二乘估计量的性质

OLS法得到如此广泛的使用，是因为它有一些理想的理论性质。高斯—马尔柯夫定理说明了这一点：若满足古典线性回归模型的基本假定，则在所有无偏估计量中， OLS估计量具有最小方差性；即OLS估计量是最优线性无偏(Best Linear Unbiased Estimator, BLUE)估计量。

我们已在第4章中讨论过最优线性无偏估计量的性质。简言之， OLS估计量b_1和b_2满足：

(1) 线性；即b_1和b_2是随机变量Y的线性函数。见式(5-16)和式(5-17)。[1]

(2) 无偏性；[2]即，

$$E(b_1)=B_1$$
$$E(b_2)=B_2$$
$$E(\hat{\sigma}^2)=\sigma^2$$

因此，平均地看，b_1和b_2将与其真实值B_1和B_2相一致，$\hat{\sigma}^2$ 将与真实的σ^2相一致。

(3) 最小方差性[3]即，

b_1的方差小于其他任何一个B_1的无偏估计量的方差。

b_2的方差小于其他任何一个B_2的无偏估计量的方差。

根据这个性质，如果我们使用 OLS法，将能够更准确地估计 B_1和B_2，虽然，其他的方法也能得到B_1和B_2的线性无偏估计量。

我们得出结论：OLS估计量有许多有用的统计性质。正因为如此，在回归分析中， OLS法才会如此广泛地应用。

蒙特卡罗试验

在理论上，OLS估计量是无偏估计量，但是，在实际中，如何知道它们是不是无偏的呢？让我们来做如下的蒙特卡罗试验(Monte Carlo Experiment)。

假定已知如下信息：

$$Y_i=B_1+B_2X_i+u_i$$
$$=1.5+2.0X_i+u_i$$

其中，$u_i \sim N(0,4)$。

即已知真实的截距和斜率分别为1.5和2.0，随机误差项服从均值为0，方差为4的正态分布，现假定X有10个观察值：1,2,3,4,5,6,7,8,9,10。

利用这些信息，可进行如下分析。利用统计软件，从 $N(0,4)$正态分布中生成10个u_i值。根

---

1 考虑，$b_2=\sum x_i y_i / \sum x_i^2$。令 $\omega_i = x_i / \sum x_i^2$ 因此，有 $b_2=\sum \omega_i y_i$ ，表明：b_2是y的线形函数，因为，y以1次方的形式出现。注意，这里我们把x_i看作非随机变量。

2 证明见Damodar N.Gujarati,*Basic Econometrics*，3d ed.,McGraw-Hill,New York，1995，pp.97-98。

3 证明见脚注6。

据给定的B_1和B_2,以及10个X值和110个生成的u_i值，通过上述等式可以得到10个Y值。称此样本为样本1。回到正态分布表，生成另外10个u_i值，得到另外10个Y值，称为样本2，按此方式，假定我们得到20多个样本。

对每个样本进行回归，得到b_1、b_2以及$\hat{\sigma}^2$。因此，可得到21个不同的b_1和b_2和$\hat{\sigma}^2$。试验及试验结果见表6-2。

表6-2 蒙特卡罗试验：$1.5+2.0X_i+u_i$；$u\sim N(0,4)$

b_1	b_2	$\hat{\sigma}^2$
2.247	1.840	2.7159
0.360	2.090	7.1663
–2.483	2.558	3.3306
0.220	2.180	2.0794
3.070	1.620	4.3932
2.570	1.830	7.1770
2.551	1.928	5.7552
0.060	2.070	3.6176
–2.170	2.537	3.4708
1.470	2.020	4.4479
2.540	1.970	2.1756
2.340	1.960	2.8291
0.775	2.050	1.5252
3.020	1.740	1.5104
0.810	1.940	4.7830
1.890	1.890	7.3658
2.760	1.820	1.8036
–0.136	2.130	1.8796
0.950	2.030	4.9908
2.960	1.840	4.5514
3.430	1.740	5.2258
$\bar{b}_1 = 1.4526$	$\bar{b}_2 = 1.9665$	$\bar{\hat{\sigma}}^2 = 4.4743$

根据表6-2给出的数据，可计算出平均的b_1和b_2及$\hat{\sigma}^2$的值分别为1.452 6，1.966 5和4.474 3,但相应的真实系数分别为1.5,2.0和4。

从这个试验我们得出什么样的结论呢？如果反复地运用最小二乘法，则平均地看，估计值将等于其真实值(总体参数)。也即OLS估计量是无偏的。在此例中，若做更多次抽样试验(多于21次)，将会得到更接近于真实值的估计值。

6.4 OLS估计量的抽样分布或概率分布

既然已经了解了如何计算OLS估计量以及其标准差，也对这些估计量的统计性质进行了研究，接下来就需要求出这些估计量的抽样分布。如果不具备上述知识，就无法辨别这些估计量接近其总体真实值的程度如何。至于估计量抽样分布的一般定义，我们已在第 3章中讨论过。(见3-2节)

为了求得OLS估计量b_1和b_2的抽样分布，我们需要在古典线性回归模型的的基本假定上再增加一条假定，即：

A6.5　在总体回归函数$Y_i=B_1+B_2X_i+u_i$中，误差项u_i服从均值为零，方差为σ^2的正态分布，即，

$$u_i \sim N(0,\ \sigma^2) \tag{6-17}$$

这个假定的理论基础是什么呢？在统计学中有一个非常著名的定理：即在第 3章中介绍过的中心极限定理(Central Limit Theorem，CLT)(参见3.2节)。

中心极限定理的内容是：独立同分布随机变量，随着变量个数的无限增加，其和的分布近似服从正态分布。

回顾第5章讨论过的误差项 u_i的特性。误差项代表了在回归模型中没有单列出来的所有其他影响因素。因为在众多的影响因素中，每种因素对 Y的影响可能都很微弱。如果所有这些影响因素都是随机的，且用 u_i代表所有这些影响因素之和，那么根据中心极限定理，能够假定误差项服从正态分布。我们已经假定了 u的均值为0，而且其方差满足同方差性假定，即为常数σ^2。因此，得到式(6-17)。

但是，u服从正态分布的假定是如何帮助我们求得 b_1和b_2的概率分布呢？这里，要用到第 3章讨论过的正态变量的另外一条性质，即正态变量的线性函数仍服从正态分布。这是否意味着：如果证明了b_1和b_2是正态变量的线性函数，那么，b_1和b_2就服从正态分布吗？你猜对了！的确可以证明这两个OLS估计量确实是正态变量u_i的线性函数。[1]

我们知道正态分布随机变量有两个特征数，均值和方差，那么，正态变量 b_1和b_2的均值和方差是什么呢？

$$b_1 \sim N(B_1, \sigma_{b_1}^2) \tag{6-18}$$

其中，

$$\sigma_{b_1}^2 = \mathrm{var}(b_1) = \frac{\sum X_i^2 \ \sigma^2}{n\sum x_i^2}$$

$$b_2 \sim N(B_2, \sigma_{b_2}^2)$$

其中，

$$\sigma_{b_2}^2 = \mathrm{var}(b_2) = \frac{\sigma^2}{\sum x_i^2}$$

简言之，b_1和b_2分别服从均值为B_1和B_2、方差为式(6-4)和式(6-6)的正态分布。

图6-4给出了这些估计量分布的几何图形。

6.5　假设检验

让我们回到widget一例，式(6-16)给出了估计的需求函数，假定价格对需求量没有影响，即，零假设为：

$$H_0: B_2=0$$

在回归分析中，这样一个“ 0 ”零假设(“ Zero ” null hypothesis)，也称之为稻草人假设(straw man hypothesis)。故意地选择这样一个假设，是为了看 Y究竟是否与X有关。如果一开始X与Y就无关，那么再检验假设，$B_2=-2$或B_2为其他任何值就没有意义了。当然，如果零假设为真，则就没有必要把 X包括到模型之中。因此，如果 X确实属于这个模型，那么，我们就期望拒绝“ 0 ”零假设H_0而接受备择假设H_1，比如说，$B_2 \neq 0$。

我们的数字结果显示：$b_2=-2.157\ 6$。因此，在widget一例中，零假设是靠不住的。但是，

1　在脚注5中，已经证明了：$b_2=\sum \omega_i y_i$。即b_2是y的线性函数。但Y本身又是u_i的线性函数，这一点可以从式(5-2)可以看出。(注：B_s和X为常量或是非随机的)。如果假定u服从正态分布，则u的线性函数Y也服从正态分布。但由于b_2是y的线性函数，因此，最终b_2是u的线性函数，所以，b_2服从正态分布。同理可证，b_1也服从正态分布。

不能仅仅看数值结果，因为由于抽样波动性，所得数值会因样本的变化而不同。显然，需要某一正式的检验过程拒绝或不拒绝零假设。我们该如何进行呢？

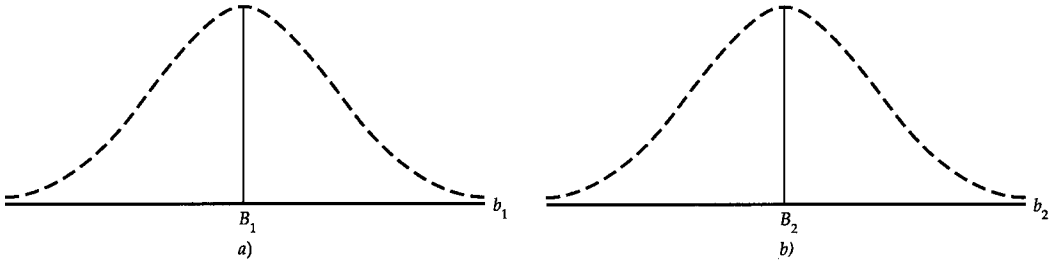


图 6-4

现在这已不成什么问题，根据式(6-19)，b_2服从均值为B_2，方差为$\sigma^2/\sum x_i^2$的正态分布。那么，根据第4章第4.5节关于假设检验的讨论，可以选择：

(1) 置信区间法

(2) 显著性检验法

对B_2和B_1的参数b_1、b_2，进行假设检验。

由于b_2服从均值为B_2，方差为$\sigma^2/\sum x_i^2$的正态分布，则变量Z服从标准正态分布，

$$Z=\frac{b_2-B_2}{se(b_2)}=\frac{b_2-B_2}{\sigma/\sqrt{\sum x_i^2}}\sim N(0,1) \tag{6-20}$$

根据第3章中介绍的标准正态分布的一条特殊性质，即正态分布的95%的区域位于($u-2\sigma$,$u+2\sigma$)之间。因此，如果零假设为$B_2=0$，计算得到的$b_2=-2.157\ 6$，我们就能够根据标准正态分布Z，求得获此b_2的概率。如果这个概率非常小，我们就能拒绝零假设，但是，如果这个概率值较大，比如说，大于10%，我们就不可以拒绝零假设。

为了用式(6-20)，必须知道真实的σ^2。而σ^2是未知的，但可以根据式(6-8)用$\hat{\sigma}^2$来估计它。如果在式(6-20)中用$\hat{\sigma}$来代替σ，则式(6-20)的右边服从自由度为($n-2$)的t分布，而不是标准正态分布，即

$$\frac{b_2-B_2}{\hat{\sigma}/\sqrt{\sum x_i^2}}\sim t_{n-2} \tag{6-21}$$

更一般地，

$$\frac{b_2-B_2}{\mathrm{se}(b_2)}\sim t_{n-2} \tag{6-22}$$

注意：在计算$\hat{\sigma}^2$时失去了两个自由度。

因此，在此情况下，为了检验零假设，需用t分布来代替(标准)正态分布，但假设检验的过程不变。

6.5.1 置信区间法

在widget一例中，共有10个观察值，因而，自由度为(10－2)=8。假定置信水平α为5%(犯第一类错误的概率)。由于备择假设是双边的，从附录A中表A-2的t分布表得：

$$P(-2.306\leq t\leq 2.306)=0.95 \tag{6-23}$$

即t值(自由度为8)位于上、下限之间的概率为95%；这个上、下限就是t的临界值。将式(6-21)代到式(6-23)，得到，

$$P\left[2.306 \leqslant \frac{b_2 - B_2}{\hat{\sigma}/\sqrt{\sum x_i^2}} \leqslant 2.306\right] = 0.95 \tag{6-24}$$

重新整理式(6-24)，得

$$P\left[b_2 - \frac{2.306\hat{\sigma}}{\sqrt{\sum x_i^2}} \leqslant B_2 \leqslant b_2 + \frac{2.306\hat{\sigma}}{\sqrt{\sum x_i^2}}\right] = 0.95 \tag{6-25}$$

或，更一般地，

$$P[b_2 - 2.306\mathrm{se}(b_2) \leqslant B_2 \leqslant b_2 + 2.306\mathrm{se}(b_2)] = 0.95 \tag{6-26}$$

式(6-26)给出了B_2的一个95%的置信区间(confidence interval)。重复应用上述过程，求得的100个这样的区间中将有95个包括真实B_2。用假设检验的语言，把这样的置信区间称为(H_0的)接受区域(region of acceptance)，把置信区间以外的区域称为(H_0的)拒绝区域(rejection region)

图6-5(a)给出了95%置信区间的几何图形。

根据第4章的讨论，如果这个区间(即接受区域)包括零假设值B_2，则不拒绝零假设。但如果零假设值落在置信区间以外(即拒绝区域)，则拒绝零假设。需时刻记住的是：无论做何种决定，都会以一定的概率(比如说5%)犯错误。

剩下的工作，就是求widget一例的具体数值了。但现在已很容易，因为我们已得到$se(b_2)$=0.120 3，将这个值代入式(6-26)，得到一个95%的置信区间，

$$-2.157\ 6 - 2.306(0.120\ 3) \leqslant B_2 \leqslant -2.157\ 6 + 2.306(0.120\ 3)$$

即，

$$-2.435\ 0 \leqslant B_2 \leqslant -1.8802 \tag{6-27}$$

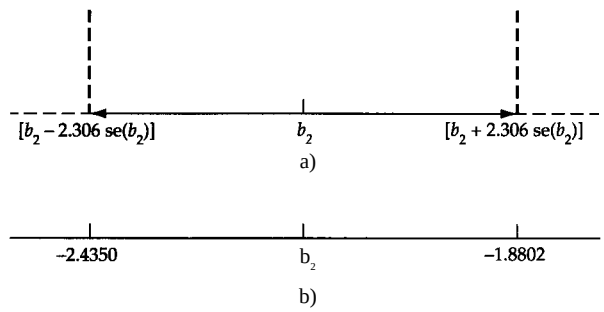


图　6-5

a)B_2的95%的置信区间　b) widget一例中的斜率的95%的置信区间

因为，这个区间没有包括零假设值0，所以拒绝零假设：价格对widget的需求量没有影响。换言之，价格确实决定了需求量。

值得注意的一点是：正如第四章中所讨论的，虽然式(6-26)式为真，但不能说，某一特定的区间[比如式(6-26)]包括真实B_2的概率为95%。因为，与式(6-26)不同，式(6-27)是固定的，而不是一个随机的区间。所以，B_2位于式(6-27)这个区间的概率为1或0。只能说如果建立像式(6-27)那样的区间100个，有95个这样的区间将包括真实B_2；我们不能保证某一特定区间一定包括B_2。

与上述过程相同，读者可以验证，截距B_1的95%的置信区间为

$$47.946 \leqslant B_1 \leqslant 51.388 \tag{6-28}$$

如果，H_0:B_1=0，H_1:$B_1 \neq 0$，显然，这个零假设将被拒绝，因为上述95%的置信区间不包括0，另一方面，如果零假设值为真，B_1=50，则不能拒绝该假设，因为有95%的置信区间包括这个

值。

6.5.2 假设检验的显著性检验法

这种假设检验方法的关键是求 t 统计量(参见第4章)及在零假设下的 t 统计量的抽样分布。决定接受还是拒绝零假设，是根据从样本数据求得的检验统计量的值而定的。

回顾 t 统计量，

$$t = \frac{b_2 - B_2}{\text{se}(b_2)}$$

它服从自由度为 $(n-2)$ 的 t 分布。如果有：

$$H_0: B_2 = B_2^*$$

其中，B_2^* 是 B_2 的某一给定值，(例如，$B_2^*=0$)，则，

$$t = \frac{b_2 - B_2^*}{\text{se}(b_2)} \tag{6-29}$$

式(6-29)很容易根据样本数据求出。由于式(6-29)右边所有的量均为已知，因此，可用计算出的 t 值作为检验统计量，它服从自由度为 $(n-2)$ 的 t 分布。与此相对应的检验过程称为 t 检验。[1]

在具体运用 t 检验时，需要知道：

(1) 对于双变量模型，自由度总为 $(n-2)$。

(2) 虽然在经验分析中常用的 α 有1%，5%或10%，但置信水平 α 是由个人任意选取。[2]

(3) 可用于单边或双边检验(参见表4-2及图4-7)。

6.5.3 Widget需求的继续

(1) 双边检验(two-tailed test)。

提出：$H_0: B_2=0$，$H_1: B_2 \neq 0$

利用式(6-29)，

$$t = \frac{-2.1576 - 0}{0.1203} = -17.94$$

根据 t 分布表，求得 t 的临界值(双边)为(见图6-6)：

显著水平	0.01	0.05	0.10
t 临界值	3.355	2.306	1.860

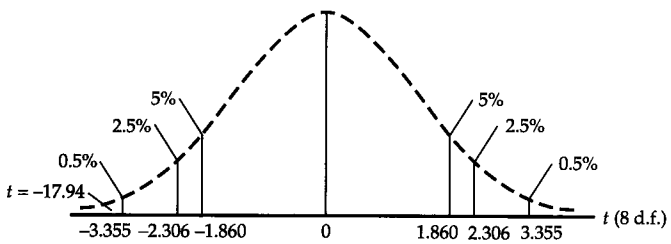


图6-6 t 分布图

1 置信区间法与显著性检验法的区别在于，前者不知道真实的 B_2 值，因而，通过建立一个 $(1-\alpha)$ 的置信区间来猜测它。而另一方面，在显著性检验方法中，我们假设真实 B_2 为某一具体值 $(=B_2^*)$，从而探求样本值 b_2 是否“充分地”按近假设值 B_2^*。

2 需要提醒注意：在第4章中有关置信水平中 p 值(即准确的置信水平)的讨论。给出检验统计量的 p 值是选择置信水平的一种好的方法。

表4-2给出了双边t检验的情况，如果计算得到的$|t|$值超过了t临界值，则将拒绝零假设。因此，在此例中，我们拒绝零假设：$B_2=0$，因为计算的$|t|$值为17.94，甚至在1%显著水平下，也远远超过了t临界值。我们得到了与置信区间法相同的结论。这并不值得奇怪，因为置信区间和显著性检验法只不过是“同一枚硬币的正反两面”。

同时，我们得到t统计量(17.94)的p值(概率值)。它甚至远远小于0.000 2的显著水平。根据t分布表，可知t值大于等于4.501的概率为0.002(自由度为8，双边检验)；由于17.94远大于4.501，因此得此t值的概率一定小于0.002。(利用计算机求解，求得获此t值(17.94)的概率约为0.000 000 5。)

(2) 单边检验(one-tailed test)

由于预期需求函数中的价格系数为负，因此，实际的假设为：$H_0:B_2=0$，$H_1:B_2<0$。

这里的备择假设是单边的。

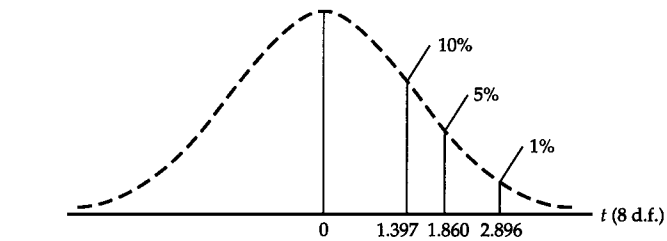
t检验过程与前面讲过的相同，只是，犯第一类错误的概率不是均等地分布在t分布的两侧，而是仅集中于一侧，或左侧或右侧。在此例中，为左侧(为什么？)。从t分布表得到t临界值(左边的)为：

显著水平	0.01	0.05	0.10
t临界值	−2.896	−1.860	−1.397

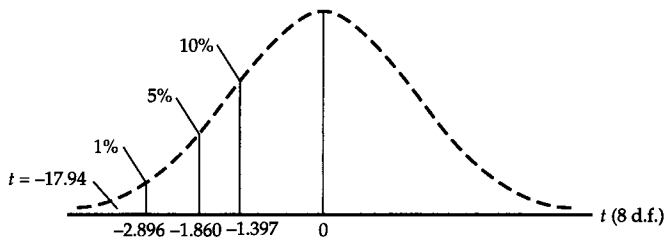
在widget一例中，在零假设$B_2=0$下，

$$t=-17.94 \tag{6-30}$$

因为在1%显著水平下，该t值比t临界值-2.896小得多，根据表4-2中的规则，我们拒绝零假设：真实的价格系数为0(参见图6-7)。[1]换言之，我们能够接受备择假设：价格系数为负。这也与经济理论相符合(参见图6-7(b))。无须赘言，我们可用同样的方法对真实的截距进行t检验。



a)



b)

图6-7　t单边检验

1　有的作者称计算的t值是高度的统计显著的。这可简单地理解为获此t值的概率极小。(前面已经看到，p值为0.000 000 5。)

6.5.4 检验σ^2的显著性：χ^2检验

回顾例3.9，样本方差S^2服从

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

在回归中，$\hat{\sigma}^2$是σ^2的估计量。因而，假定随机样本来自正态总体，可以证明，

$$\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)} \tag{6-31}$$

它服从自由度为$(n-2)$的χ^2分布。因此，如果在零假设下，给定σ^2的值，就能用χ^2检验来检验真实的σ^2。我们仍用widget一例来说明。

在widget一例中，$\hat{\sigma}^2$=1.1939。假设真实的σ^2为1.5。我们能接受这个假设吗？利用式(6-31)，求得

$$\frac{8(1.193\ 9)}{1.5} = 6.367\ 5 \tag{6-32}$$

根据附录A-4的χ^2分布表，查得χ^2值大于或等于6.37的概率在0.75~0.50之间。(自由度为8)，这比通常用的0.01或0.05略高，借助计算机，我们求得获此χ^2值(6-37)的p值约为0.61。因而，若在显著水平α为0.05或0.10下，我们不能拒绝假设：真实σ^2为1.5。换言之，样本值1.939与真实值1.5不是有显著差别的。

上述讨论表明，一旦知道给定估计量所服从的概率分布(或抽样分布)，那么，对这一估计量的真实值进行假设检验将会很容易。

6.6 拟合优度的检验：判定系数r^2

上一节中的分析表明，根据t检验，估计的斜率和截距均为统计显著的，这说明式(6-16)样本回归函数很好地拟合了数据。当然，并非每一个Y值都准确落在了估计的总体回归线上。也就是说，并非所有的$e_i = Y_i - \hat{Y}_i$都为0；从表5-4可以看到，有些e值为正，有些则为负。能否建立一个“拟合优度”的度量规则，从而辨别估计的回归线拟合真实Y值的优劣？的确可以，我们称之为判定系数(coefficient of determination)，用符号r^2表示。下面来看如何计算r^2？

回顾，

$$Y_i = \hat{Y}_i + e_i \tag{5-5}$$

将式(5-5)稍微变形，得到：

$$(Y_i - \bar{Y}) = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \tag{6-33}$$

(Y_i与其均值的偏差)　($\hat{Y}_i$与其均值的偏差)　(即e_i)(残差)

用小写字母表示与均值的偏差，得到：

$$y_i = \hat{y}_i + e_i \tag{6-34}$$

(注：$y_i = Y_i - \bar{Y}, \bar{e} = 0$，所以，$\bar{\hat{Y}} = \bar{Y}$，即真实$Y$值的均值等于估计$Y$值的均值。)

或，

$$y_i = b_2 x_i + e_i \tag{6-35}$$[1]

现对(6-34)式两边同时平方再求和，经过数学变形后，得到：

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 \tag{6-36}$$

或等价地，

$$\sum y_i^2 = b_2^2 \sum x_i^2 + \sum e_i^2 \tag{6-37}$$[2]

---

1 SRF：$Y_i = b_1 + b_2 X_i + e_i$，用偏差形式(即每一个变量表示为与其样本均值的偏差)变为$y_i = b_2 x_i + e_i, \hat{y}_i = b_2 x_i$。用偏差的形式，截距$b_2$并在模型中出现。

2 对式(6-34)两边平方再求和，得到

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 + 2\sum \hat{y}_i e_i$$
$$= b_2^2 \sum x_i^2 + \sum e_i^2$$

由于，$\sum \hat{y}_i e_i = b_2, \sum x_i e_i = 0$。注：$x_i = X_i - \bar{X}$

注：$\hat{y}_1 = b_2 x_i$

随后将会看到，这是一个重要的关系。

式(6-36)中出现的各种平方和定义如下：

$\sum y_i^2$ 表示总离差平方和(total sum of squares,TSS)，[1]

$\sum \hat{y}_i^2$ 表示回归平方和(explained sum of squares, ESS)，

$\sum e_i^2$ 表示残差平方和(residual sum of squares, RSS)。

则式(6-36)可简化成为，

$$\text{TSS}=\text{ESS}+\text{RSS} \tag{6-38}$$

式(6-38)表明：Y值与其均值的总离差可以分解为两个部分：一部分归于回归线，另一部分归于随机因素，因为并不是所有的真实观察值 Y都落在拟合的直线上，我们可以从图 6-8中清楚地看到这一点。

现在，若要选择好的拟合样本回归函数，则要求 ESS比RSS大得多：若所有真实的 Y值都落在拟合的样本回归线上，则 RSS将为0；另一方面，如果样本回归线拟合的不好，则 RSS将比ESS大得多；这里，存在的极端的情形是，如果 X不能解释 Y，则ESS将为0，而RSS等于TSS。但也存在相反的情况。一般典型地情况是： ESS和RSS均不为零，如果ESS相对地比RSS大，则样本回归函数将能在很大程度上解释 Y的变动；如果RSS比ESS相对大一些，则样本回归函数只能部分解释Y的变动。所有这些定性分析，从直觉上易于理解，但是，能否从定量的角度来分析呢？如果把式(6-38)两边同除以TSS，得到，

$$1 = \frac{\text{ESS}}{\text{TSS}} + \frac{\text{RSS}}{\text{TSS}} \tag{6-39}$$

定义，

$$r^2 = \frac{\text{ESS}}{\text{TSS}} \tag{6-40}$$

我们称r^2为判定系数，通常用来度量回归线的拟合优度。用文字表述即为，判定系数度量了回归模型对Y的变动解释的比例(或百分比)。

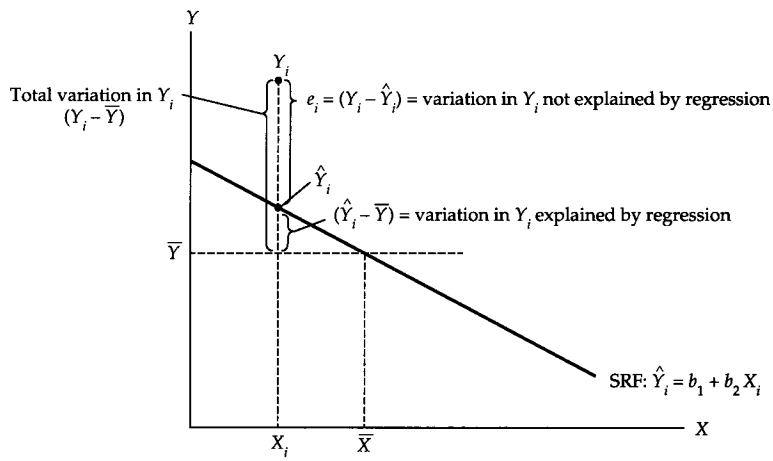


图6-8　Y_i中总变动的分解

1　术语离差与方差不同。离差表示变量与其偏差的平方和。方差表示离差除以相应的自由度。简言之，方差=离差/自由度。

下面给出r^2的两条性质：

(1) 非负性(为什么？)。

(2) $0 \leq r^2 \leq 1$，因为部分(ESS)不可能大于整体(TSS)，[1]若$r^2=1$，表示线性模型完全解释Y_i的变动，若$r^2=0$，则表示Y与X之间无任何关系。

6.6.1 r^2的计算公式

根据式(6-40)，式(6-39)可改写为

$$1 = r^2 + \frac{\text{RSS}}{\text{TSS}}$$

$$= r^2 + \frac{\sum e_i^2}{\sum y_i^2} \tag{6-41}$$

因此，

$$r^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} \tag{6-42}$$

还有几个等价的公式计算r^2，参见习题6.5。

6.6.2 Widget一例中的r^2

根据表5-4中的数据，利用式(6-42)，我们得到Widget一例中的r^2值：

$$r^2 = 1 - \frac{9.551\,5}{393.60} \tag{6-43}$$

$$= 0.975\,7$$

因为，r^2最大为1，所以，计算的r^2值已经相当大了。在需求函数中，价格变量X能以98%的程度解释需求量的变动。因此，在此例中，可以认为样本回归函数(6-16)相当好地拟合了总体回归函数。

6.6.3 相关系数r

在第2章2.8节中，我们介绍了样本相关系数(sample coefficient of correlation)r，它是度量两变量X与Y之间线性相关程度的指标，r可由式(2-53)计算得到，也可写为

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \tag{6-44}$$

$$= \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}} \tag{6-45}$$

但是，相关系数也能够根据判定系数r^2来计算：

$$r = \pm\sqrt{r^2} \tag{6-46}$$

许多回归软件都有计算r^2程序,因此很容易求得r。惟一的问题是关于r的符号。但根据问题的实际意义很容易决定r的符号。在widget一例中,由于预期需求量与价格负相关,因此,r值应为负。一般而言，r与斜率同号,我们可从式(5-17)和式(6-45)中清楚地看到这一点。因此，在Widget一例中，

$$r = -\sqrt{0.9757} = -0.987\,8 \tag{6-47}$$

根据表5-4数据并利用式(6-45)，读者很容易验证上式给出的r值。在2.8节中我们已经讨论了r的性质以及r可能的几种图形(见图2-7)。

1 假定回归模型中包括截距项。更详细的讨论见习题6.15。

在回归分析中，判定系数r^2比相关系数r更有意义。判定系数告诉我们解释变量对应变量变化的解释程度，因而它全面地度量了一个变量决定另一个变量变动的程度。但是，r却不能。除此之外，我们还将看到，在多元回归分析中，对r的解释也是不确定的。

6.7 回归分析结果的报告

有不同种的方式报告回归分析的结果，但在本书，我们将沿用如下形式，仍以widget一例为例：

$$\begin{aligned} &\hat{Y}_t = 49.667\,0 - 2.157\,6\,X_t && r^2 = 0.9757 \\ &\text{se} = (0.746\,4)\quad (0.120\,3) && \text{d.f.} = 8 \\ &t = (66.542)\quad (-17.935) && p\text{值} = (0.000)^*\ (0.000)^* \end{aligned} \tag{6-48}$$

*表示非常小

在方程(6-48)中，第一组括号内的数表示估计的回归系数的标准差，第二组括号内的数表示在零假设：每个回归系数的真实值为零下，根据式(6-22)估计的t值的t值。(即t值是估计的系数与其标准差之比)。第三组括号内的数表示(t值的)p值。[1]从现在起，如果没有特别地规定零假设，习惯地假定为“零”零假设(即假定总体参数为零)。如果拒绝该零假设(即检验统计量是显著的)，则表示真实的总体参数值不为零。

用上述形式报告回归结果的一个优点是，可以一目了然地看到每一个估计的系数是否是统计独立的，即是否显著不为零。通过列出p值，能够决定估计的t值的精确的显著水平。因此，估计的斜率的t值为－17.935，其p值几乎为0。这意味着如果零假设为真，即真实的价格系数为零，则得此t值为17.935(绝对值)的概率几乎为0。因而拒绝零假设。正如我们第四章中所讲的，p值越低，拒绝零假设的证据就越充分。

p值的一个优点在于无须规定显著水平α在某一任意的水平，比如1%，5%或10%。在Widget一例中，实际的p值为0.000 000 5。在此显著水平下，如果拒绝零假设，即价格系数为零，那么，犯错误的机会为0.000 005。

当然，任何其他的零假设(除了“零”零假设以外)都很容易用前面讨论过的t检验来检验。因此，如果零假设为真实的截距为49，备择假设$H_1 \neq 49$，则t值为：

$$t = \frac{49.667\,0 - 48}{0.746\,4} = 0.893$$

读者很容易验证，在5%的显著水平下，根据双边t检验(临界值$t_{0.025,8}$=2.306)，这个t值是统计不显著的。前面已经讲过，“零”零假设是最基本的一种稻草人假设。

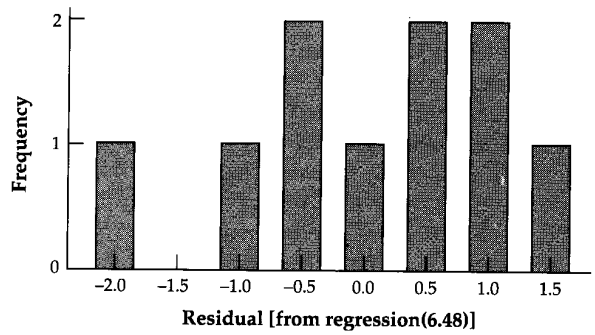


图 6-9

1 本书末尾给出的t分布表能够用电子表(对若干数字计算出p值)来代替。对正态分布，χ^2分布以及F分布表也是如此。

6.8 正态性检验

在结束Widget一例之前，我们需要看一下式(6-48)的回归结果。别忘了该统计检验过程是建立在假定误差项 u_i 服从正态分布的基础之上的。既然我们不能直接地观察真实的误差项 u_i，那么如何从本例中证实 u_i 确实服从正态分布呢？我们有 u_i 的近似值 —— 残差 e_i，因此，可通过 e_i 来获悉 u_i 的正态性。对正态性的检验方法有若干种，这里，仅介绍三种相对简单的检验方法。[1]

6.8.1 残差直方图

残差直方图是用于获悉随机变量的概率密度函数(PDF)的形状如何。在横轴上，我们将变量值(即OLS的残差)划分为若干适当的区间，在每一个区间，建立高度与观察值个数(即频率)相一致的长方形。在Widget一例中，得到如图6-9所示的直方图。如果把钟型正态曲线叠加在直方图上，你就会对所关注变量概率分布的性质有所了解。在Widget一例中，由于我们仅有10个观察值，因而随机误差项的残差直方图近似正态分布的图形。

一种好的实践方法就是作出回归残差直方图，这样可以粗略地了解其概率分布的形状。

6.8.2 正态概率图

另一种研究随机变量概率密度函数相对简单的图形方法是正态概率图(Norma Probability Plot, NPP)(在专用的正态概率纸上作图)。在横轴上(X轴),标出所关注变量的值，在纵轴上(Y轴),标出该变量服从正态分布所对应的均值。因此，若该变量的确来自正态总体，则正态概率图将近似为一直线。图6-10给出了利用MINTAB软件作出Widget一例的正态概率图(残差)。

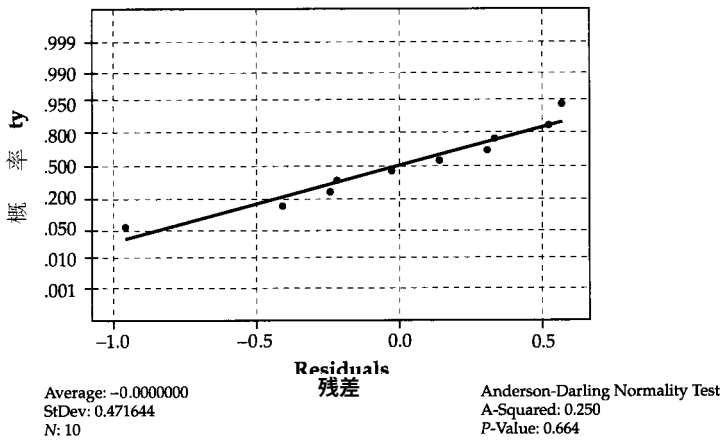


图6-10 Widget回归方程的残差

Minitab计算出每一观察值发生的概率……(假定是正态分布)并将所得概率值取对数。纵坐标是概率值，横坐标是数据值。最小二乘直线拟合了图中的数据点。该直线估计了总体累积分布函数。图中给出了总体参数 u 和 σ 的估计值(MINTAB参考手册，第1-26页)。

前面已经提到，如果在正态概率图中拟合线近似为一条直线，那么，我们所关注的变量服从正态分布。在图6-10中，Widget回归方程的残差就近似正态分布，因为这条直线看似很好地拟合了数据点。

MINTAB也可作Anderson-Darling正态检验，称之为 A^2 统计量(A^2 statistic)。该正态检验假

1 对检验正态性的各种方法的详细讨论参见 G.Barrie Wdtherill, *Regression Analysis with Applications*, Chapman and Hall, London,1986，Chapter 8。

设变量服从正态分布。从图 6-10可以看出，在Widget一例中，计算的 A^2统计量为0.250，获此 A^2值的概率为0.664(相当大)，因此，我们不能拒绝零假设，即残差服从正态分布。从图 6-10中还得知该(正态)分布的均值为 0，标准差约为0.47。

6.8.3 Jarque－Bera检验

目前，另一种常用的正态性检验是 Jarqe-Bera(JB)检验，[1]在许多统计软件中也都包括这种检验方法。它是依据 OLS残差，对大样本的一种检验方法(或称为渐近检验)。首先计算偏度系数S(对概率密度函数对称性的度量)及峰度系数 K(对概率密度函数的“ 胖瘦 ”的度量)。(参见第2章)；对于正态分布变量，偏度为 0，峰度为3，(见图2-8)

Jarqe和Bera建立了如下检验统计量：

$$JB = \frac{n}{6}\left[S^2 + \frac{(K-3)^2}{4}\right] \tag{6-49}$$

其中，n为样本容量，S为偏度，K为峰度。

他们证明了：在正态性假定下，式(6-49)给出的JB统计量渐近地服从自由度为 2的χ^2分布，用符号表示为，

$$JB_{asy} \sim \chi^2_{(2)} \tag{6-50}$$

其中，asy表示渐近地。

从式(6-49)可以看出，如果变量服从正态分布，则 S为0，K为3，因而JB统计量的值为零。但是如果变量不是正态变量，则 JB统计量将为一个逐渐增大值。我们很容易从 χ^2分布表计算出JB统计量的值。在某一显著水平下，根据式(6-49)计算的值超过临界的 χ^2值，则将拒绝正态分布的零假设；但如果没有超过临界的 χ^2值，则不能拒绝零假设。当然，若计算出 χ^2值的p值，则将知道得此χ^2值精确的概率。

由于在Widget一例中，仅有 10个观察值，严格地说， JB检验是无效的。但若机械地运用JB检验，则得到JB统计量的值为0.516 9(p值为0.772 2)。这表明计算得到的JB统计量不是统计显著的。因此，我们不能拒绝零假设： widget回归的残差服从正态分布。当然，这个结论并不是很准确，因为样本容量太小了。

同时我们得到Widget一例中的偏度系数为－ 0.461 4，表明残差是负偏的，(参见图 6-9)；峰度系数为2.376 3，表明残差的分布比正态分布略胖。[2]

总之，在实际中，在古典回归模型假定下，用上述一两种方法对正态性进行检验是很重要的。因为假设检验的过程在很大程度上依赖于正态性这个假定，尤其是当样本容量很小时。

6.9 关于计算：回归分析的软件

在本章和前面的几章中，给出了若干计算公式。我们可用已过时的计算器来计算各种公式。但是，计算机为我们提供了极大的方便。今天 ,不仅有许多好的商业回归软件，而且许多读者还已经拥有了计算机，当然这要归功于个人电脑的普及。因此，在计算机时代，人们很少再在计算器来做如此复杂的工作了。或许你会用它来做一些简单的回归习题，不过那也只是为了熟悉本章所出现的各种公式罢了。许多回归软件中已能计算这些公式。目前，较为流行的软件有ET，EVIEWS,LIMDEP,SHAZAM,TSP,SAS,SPSS以及MINTAB。本书的大部分实例都是用上述

1 参见C.M.Jarque，A.K.Bear,“ A Test for Normality of Observations and Regression Residuals ”, *International Statistical Review*，vol.55，1987，pp.163-172。

2 有关偏度和峰度的详细讨论见本书第 2章。

的一些软件来分析的。

6.10 实例：美国进口支出

我们将通过一个建立在具体经济数据之上的实例，来对上面讨论的双变量模型做一概括。表6-3给出了1968~1987年间美国的进口支出(或外国商品，包括耐用的或非耐用的)及个人可支配收入(PDI)(税后收入)的数据。以1982年的美元价为基准度量单位。用恒定的美元值，以消除通货膨胀的影响。

凯恩斯著名的消费函数理论表明：个人的消费支出(PCE)与PDI正相关。由于对外国物品的消费是总消费支出的一部分，因此，可以认为进口支出与PDI正相关。模型的形式为：

$$Y_t = B_1 + B_2 X_t + \mu_t \tag{6-51}$$

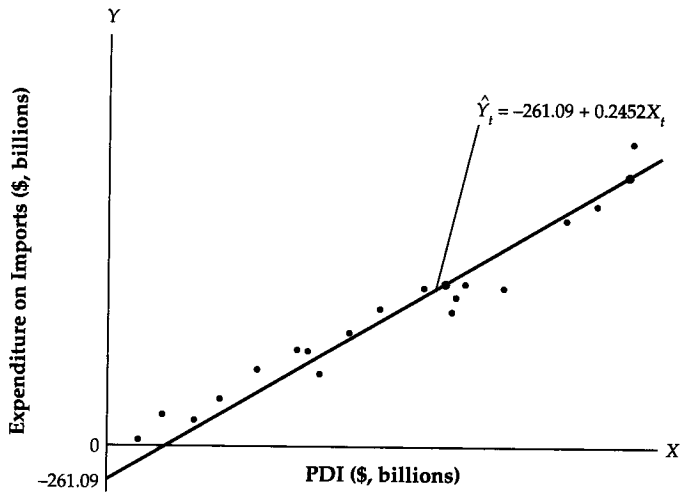
(其中，Y表示进口的消费支出，X表示PDI，下标t表示样本数据为一时间序列)拟合表6-3给出的数据，看一看这两个变量之间是否呈线性关系。对X、Y作散点图，见图6-11。

表6-3 U.S进口商品支出(Y)与个人可支配收入(X)，1968~1987[①]

年	Y	X	年	Y	X
资料来源：总统经济报告，1989。数据Y见表B-21，第331页，数据X见表B-27，第333页。					
1968	135.7	1551.3	1978	274.1	2167.4
1969	144.6	1599.8	1979	277.9	2212.6
1970	150.9	1668.1	1980	253.6	2214.3
1971	166.2	1728.4	1981	258.7	2248.6
1972	190.7	1797.4	1982	249.5	2261.5
1973	218.2	1916.3	1983	282.2	2331.9
1974	211.8	1896.9	1984	351.1	2469.8
1975	187.9	1931.7	1985	367.9	2542.8
1976	229.9	2001.0	1986	412.3	2640.9
1977	259.4	2066.6	1987	439.0	2686.3

① 单位为亿美元(1982)。

图 6-11



散点图表明：两变量之间的关系近似为线性关系，因此，线性模型(6-21)可能是适当的。在随后的一章中，我们将讨论：如果模型违背了古典线性回归模型的某些假定，我们将对模型重新讨论。现在，继续我们的分析！

在求OLS结果之前，先看一下经济理论是如何解释 B_1、B_2的。B_2(回归)数学的意义表示直线的斜率，但从经济学的角度讲，它表示出口商品的边际消费倾向(Marginal Propensity to Spend, MPS)，即PDI每增加1美元所引起的出口商品平均支出的变化量。例如， MPS为0.25，即如果PDI每增加1美元，则对出口商品的平均消费支出将增加 25美分。经济理论表明MPS应为正，但小于1，即$0<B_2<1$。这是因为，随着收入的增加，我们不会将增加的所有收入都用于消费，其中一部分将用作储蓄。

截距B_1又有何意义呢？我们知道，截距是当 X取0值时的Y值。在此例中，它表示 PDI为零时，进口商品的平均消费支出。一般来说，在此情况下的消费支出也应该为零(为什么？)。然而不幸的是，估计出的截距往往不为零。因此，虽然在某些情况下，可给出截距特殊的经济意义(当遇到这类情况时，我们会对截距谈得更多一些。)，但是，一般地，它都没有什么具体的经济意义，只不过是一个数字实体。

回归所需的原始数据在习题 6.16中给出，这里我们用 SHAZAM统计软件进行计算，表 6-4给出了实际的计算结果。

表6-4　进口支出/PDI回归结果(SHAZAM)

(续)

OLS ESTIMATION
20 OBSERVATIONS　　DEPENDENT VARIABLE = Y
...NOTE..SAMPLE RANGE SET TO:　　1,　20

R-SQUARE = 0.9388　　R-SQUARE ADJUSTED = 0.9354
VARIANCE OF THE ESTIMATE = 475.48 = $\hat{\sigma}^2$
STANDARD ERROR OF THE ESTIMATE = 21.806 = $\sqrt{\hat{\sigma}^2}$

MEAN OF DEPENDENT VARIABLE = 253.08 ($\overline{Y}$)

ANALYSIS OF VARIANCE – FROM MEAN

	SS	DF	MS	F	See Chapter 7
REGRESSION	0.13127E+06	1.	0.13127E+06	276.083	
ERROR	8558.7	18.	475.48		
TOTAL	0.13983E+06	19.	7359.6		

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 18 DF	① PARTIAL CORR.	② STANDARDIZED COEFFICIENT	③ ELASTICITY AT MEANS
X	0.24523	0.14759E-01	16.616	0.9689	0.96891	2.0317
CONSTANT	−261.09	31.327	−8.3345	−0.8912	0.00000E+00	−1.03

OBS. NO.	OBSERVED VALUE (Y_i)	PREDICTED VALUE ($\hat{Y}_i$)	CALCULATED RESIDUAL (= e_i)	⑦
1	135.70	119.34	16.364	|　*
2	144.60	131.23	13.370	|　*
3	150.90	147.98	2.9212	|*
4	166.20	162.77	3.4338	|*
5	190.70	179.69	11.013	|　*
6	218.20	208.85	9.3548	|　*
7	211.80	204.09	7.7123	|　*
8	187.90	212.62	−24.722	*　|
9	229.90	229.62	0.28372	*
10	259.40	245.70	13.697	|　*
11	274.10	270.42	3.6772	|*
12	277.90	281.51	−3.6072	*|
13	253.60	281.92	−28.324	*　|
14	258.70	290.34	−31.636	*　|
15	249.50	293.50	−43.999	*　|
16	282.20	310.76	−28.563	*　|
17	351.10	344.58	6.5193	|*
18	367.90	362.48	5.4174	|*
19	412.30	386.54	25.760	|　　*
20	439.00	397.67	41.327	|　　*

按照式(6-48)的形式，得到如下回归结果：

④DURBIN-WATSON = 0.5951 ④VON NEUMAN RATIO = 0.6264 ④RHO = −0.73339
RESIDUAL SUM = −0.17764E−12 RESIDUAL VARIANCE = 475.48 ($\hat{\sigma}^2$)
SUM OF ABSOLUTE ERRORS= 321.70 = $\sum|e_i|$
⑤R-SQUARE BETWEEN OBSERVED AND PREDICTED = 0.9388 (= R^2)
⑥RUNS TEST: 5 RUNS, 14 POSITIVE, 6 NEGATIVE, NORMAL STATISTIC = −2.4326

Source: SHAZAM output based on Table 6-3.
Notes: ① See Chap. 10.

② Obtained by regressing $[(Y_i - \bar{Y})/s_y]$ on $[(X_i - \bar{X})/s_x]$. These are called standardized variables. In such a regression the intercept value is zero.

③ Elasticity = slope $(\bar{X}/\bar{Y})$ where $\bar{X}$ and $\bar{Y}$ are means of X and Y.
④ Discussed in Chap. 12.
⑤ See problem 6.5.
⑥ See Chap. 12.
⑦ Residual graph (see Chaps. 11 and 12).

$$\hat{Y}_t = -261.09 + 0.245\,23X_t \qquad r^2 = 0.938\,8$$
$$\text{se} = (31.327) \quad (0.014\,8) \qquad \text{d.f.} = 18 \qquad (6\text{-}52)$$
$$t = (-8.334) \quad (16.599\,6)$$
$$p \text{ 值} = (0.000)^* \quad (0.000)^*$$

注 *表示值很小。

回归直线见图6-11。

6.10.1 对回归结果的解释

回归方程(6-52)以及图6-11表明：与预期相同，进口支出与PDI正相关。斜率0.2452表示进口商品的边际消费倾向为24美分，即PDZ每增加1美元，平均而言，对进口商品的消费支出将增加24美分。与预期相同，MPS也为正且小于1。

截距−261表示当PDI为零时，平均支出为负的261美元。我们已经强调过，截距可能没有什么明显的经济意义，本例亦如此。

r^2值为0.9388，意味着模型能够以94%的比例解释进口产品支出的变动。也就是说，仅通过一个解释变量PDI就能很好解释应变量的变化。由于r^2最大为1，因此，观察到的r^2值已经很大了，表明：所选择的模型很好地拟合了实际数据。

6.10.2 回归结果的显著性检验

估计的回归系数是统计显著的吗？也就是说，回归系数显著不为零吗？我们立刻可从估计的回归式(6-52)中得到答案。首先，来看斜率。在零假设：真实斜率为零下，计算得到的t值为16.616。(参见6.7节)。我们拒绝零假设吗？正如前面讨论的那样，(对于某一给定置信水平)如果计算得到的t值超过t的临界值,那么就拒绝假设。在本例子中，自由度为18(为什么？)，若选择置信水平α=5%，并有备择假设H_1—— 真实的B_2大于零。(为什么这样假设？)—— 则右侧的t临界值为

$$t_{0.05,18} = 1.743$$

由于计算得到的t值16.62远大于t临界值，根据t检验，能够拒绝零假设，即PDI对进口商品支出无影响。(得此t值的准确p值为$1.125/10^{12}$)。若用置信区间法(confidence interval approach)，可得到相同的结论。根据式(6-26)建立区间，

$$p[b_2 - t_{\alpha/2,df}\text{se}(b_2) \le B_2 \le b_2 + t_{\alpha/2,df}\text{se}(b_2)] = (1-\alpha) \qquad (6\text{-}53)$$

如果α=5%，则$t_{\alpha/2,18} = 2.101$，式(6-53)变为，

$$[0.245\ 23-2.101(0.014\ 759)]\leq B_2\leq[0.245\ 3+2.101(0.014\ 59)]$$

即为，

$$0.241\ 2\leq B_2\leq 0.276\ 2 \tag{6-54}$$

即为B_2的95%的置信区间。显然，95%的置信区间(或是接受区域)不包括零假设值：$B_2=0$，它位于拒绝区域内。因而，拒绝该零假设。

简言之，用置信区间法和t检验法得到相同的结论。这并没有什么可奇怪的(为什么？)。

留给读者证明：用置信区间法或t检验法得到截距也是统计显著的，虽说本例的截距并没有什么经济含义。读者还需作出其正态概率图看一看误差项是否服从正态分布。

6.11　预测

根据1968～1987年的历史数据，我们得到了式(6-52)进口需求函数(确切地说，是对进口商品的消费支出)，能用它来进行预测吗？也就是说，给定一个PDI值，能预测平均地消费支出吗？(称之为均值预值)答案是肯定的。

假定变量X(PDI)取某一值X_0，(比如1988年的X_0=2 800美元)现在估计1988年的$E(Y|X_0)$，即相应PDI的平均消费支出。

现令

$$\hat{Y}_0=E(Y\mid X_0)\text{的估计量}$$

如何得到这个估计值呢？在CLRM假定下，将相应的X值代入式(6-52)得到，

$$\begin{aligned}\hat{Y}_{1988}&=-261.09+0.245\ 23(2\ 800)\\&=425.556\end{aligned} \tag{6-56}$$

即预测的1988年对进口产品的平均支出为426亿美元(1982)。

虽然经济计量理论表明在CLRM的假定下，$\hat{Y}_{1988}$或更一般的$\hat{Y}_0$是真实均值的无偏估计量(也即总体回归线的一个点)，但对任一给定样本，$\hat{Y}_0$不可能等于其真实均值(为什么？)。两者之差称为预测误差。为了估计这个误差，需要求出$\hat{Y}_0$的抽样分布。[1]在CLRM假定下，可以证明$\hat{Y}_0$服从正态分布，其均值，方差分别为：

$$\text{均值}=E(Y\mid X_0)=B_1+B_2X_0 \tag{6-57}$$

$$\text{方差}=\sigma^2\left[\frac{1}{n}+\frac{(X_0-\bar{X})}{\sum x_i^2}\right] \tag{6-58}$$

其中，$\bar{X}$表示回归函数式(6-52)中X的样本均值，

$\sum x_i^2$表示X的真实值与均值的偏差平方和，

σ^2表示u_i的方差，

n表示样本容量。

式(6-58)的正平方根为$\hat{Y}_0$的标准差，se($\hat{Y}_0$)。

在实际中，σ^2未知，若用其无偏估计量$\hat{\sigma}^2$代替，则$\hat{Y}_0$服从自由度为$(n-2)$的t分布，(为什么？)因此，我们可用t分布对与X_0相关的真实(即总体)均值Y建立一个$100(1-\alpha)$%的置信区间：

$$\begin{aligned}\text{var}\,\hat{Y}_{1988}&=475.48\left[\frac{1}{20}+\frac{(2\ 800-2\ 096.7)}{2\ 182\ 851}\right]\\&=131.5243\end{aligned} \tag{6-60}$$

1　注意$\hat{Y}_0$是一个估计量，因此他有抽样分布。

因此，

$$\mathbf{se}(\hat{Y}_0) = \sqrt{131.5243} = 11.468\ 40 \tag{6-61}$$

注：$\bar{X}$ =2 096.7，$\sum x_i^2$ = 2 182 851,$\hat{\sigma}^2$ = 475.48

上述结果表明：若PDI的估计值为2800亿美元，则预测的平均消费支出为425.47亿美元。这个预测值的标准差约为11.4亿美元。

现在假定我们要对1988年总体平均的消费支出建立一个95%的置信区间。根据式(6-59)得：

$$425.556 - 2.101(11.468\ 4) \leq E(Y \mid X_0 = 2\ 800) \leq 425.556 + 2.101(11.468\ 4)$$

即，

$$401.460\ 9 \leq E(Y \mid X_0 = 2\ 800) \leq 449.651 \tag{6-62}$$

注：自由度为18，在5%显著水平上，双边的t值为2.101

若1988年，PDI的值为2 800亿美元，式(6-62)表明：虽然1998年的平均进口消费支出惟一的一个最优估计值为425.56亿美元，但预期该值以95%的置信度，落在区间401.46亿美元至449.65亿美元之间。

如果对每一个X值，我们得到诸如式(6-22)一个95%的置信区间，则可以得到对应于每个X，也就是对应于整条总体回归线的真实平均消费支出的置信区间或者说置信带，从图6-12中可以很容易看到这一点。

从图6-12中，还可发现其他一些有趣的事实。当$X_0 = \bar{X}$时，置信带的宽度最小，这一点很容易从式(6-58)方差的计算式中得以证明。但是，随着X_0逐渐远离$\bar{X}$，置信带将迅速加宽(即预测的误差将会增加)，这表明：随着X_0逐渐远离$\bar{X}$，历史回归的预测能力将显著地减弱。因此，在用“外推”历史回归线来预测Y的均值时，需要格外谨慎。就实践而言，不宜用式(6-52)的进口支出回归方程来预测2001年的平均进口支出。

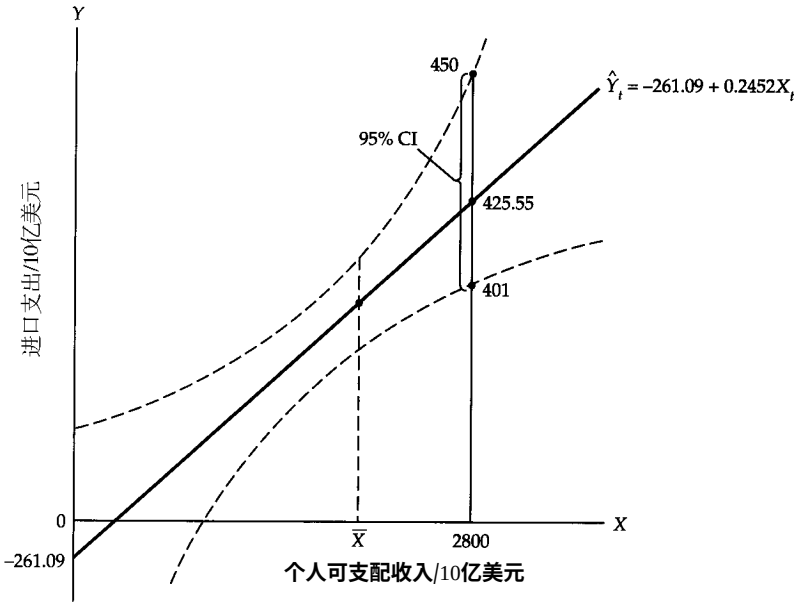


图6-12 进口支出均值的95%的置信区间

6.12 实例

名义汇率与相对价格的关系

回到例5.5。式(6-48)完整的回归结果如下：

$$\begin{aligned} \hat{Y}_t &= 6.682 - 4.318X_t & r^2 &= 0.582\ 8 \\ \text{se} &= (1.222)\quad (1.133) & \text{d.f.} &= 13 \\ t &= (5.47)\quad (-3.81) & & \\ p\text{值} &= 0.000\ 1\quad 0.002 & & \end{aligned} \tag{6-63}$$

从上面的回归结果可以看出：截距和斜率都显著不为零，因为 p值都很小。比如，与变量 X_t相关的t系数的p值为0.002，表示得到t值等于或小于－3.81的概率仅为0.002。如果零假设：真实的斜率为零为真，那么，计算的斜率值应该与零值差别不大，而在此情况下的 t值会接近零。[注：$t=(b_2-B_2)/\text{se}(b_2)$]。但是事实并非如此，计算的 B_2与零值相差甚远；事实上，它与零值相差大约3.81个标准差。[1]因此，我们拒绝零假设：真实的 B_2为零。在拒绝零假设的过程中，我们犯第一类错误的概率非常小，约为0.002。

r^2值为0.528，表示两个国家的相对价格解释了马克对美元汇率变动的 53%。由于r^2的最大值为1，所以这个r^2值比较低。我们会担心“较低”的r^2值吗？不会。原因留到下一章解释。

6.12.1 人均消费支出与人均可支配收入

表6-5给出了美国1980～1995年间人均消费支出(Per capita Consumption Expecnditure, PCE)和人均可支配收入(Per capita Disposable Personal Income, PDPI)的数据。为了观察PCE与PDPI间的关系，得到下面OLS回归结果：

$$\begin{aligned} \widehat{PCE} &= -3\ 116 + 1.095\ 1\text{PDPI} & & \\ \text{se} &= (453.5)\quad (0.026\ 6) & r^2 &= 0.992 \\ t &= (-6.87)\quad (41.08) & & \\ p &= (0.000^*)\quad (0.000^*) & & \end{aligned}$$

*表示值很小。

与预期相同，人均消费支出与人均可支配收入正相关。斜率可解释为边际消费倾向(MPC)——每增加一美元的可支配收入所增加的平均消费支出。由于p值很小，所以斜率是高度统计显著的。值得惊讶的是这个MPC明显比1大，因为，

$$t = \frac{(1.095\ 1 - 1.0)}{0.026\ 6} = 3.575\ 2$$

在0.5%的显著水平下(自由度为14)，t值是显著的(单边检验)。根据经济理论，MPC小于1。为什么本例中的MPC会大于1呢？的确这是一个很有意思的问题。

表6-5 美国1980 ~ 1995 PCI 与PDPI(美元，1992)

年份	PCE'92	PDPI'92
1980	13216	14813
1981	13245	15009
1982	13270	14999
1983	13829	15277
1984	14415	16252
1985	14954	16597
1986	15409	16981
1987	15740	17106
1988	16211	17621
1989	16430	17801
1990	16532	17941
1991	16249	17756
1992	16520	18062
1993	16809	18078
1994	17159	18330
1995	17400	18799

资料来源：总统经济报告，1997，表B-29，第333页。

6.12.2 MBA工资的回顾

参考例5.6。完整的回归结果如下：

1 在零假设$B_2=0$下，t值变为$b_2/\text{se}(b_2)$。因此，这时的t值表明了b_2偏离零值的方差有多少。事实上，$|t|$值越大，拒绝零假设的证据就越充分。

$$\widehat{ASP}_i = -33\,598 + 648.08\text{GMAT}$$
$$\text{se}=(45\,702) \quad (73.16)$$
$$t=(-7.35) \quad (8.86)$$
$$p=(0.000)^* \quad (0.000)^* \qquad r^2=0.737 \qquad (6\text{-}65)$$

*表示值很小。

读者很容易验证斜率显著不为零，因为计算所得的 p 值很小。或许，要得到一个高的GMAT分数，的确需要付出代价。

6.13 小结

上一章我们讨论了如何对双变量线性回归模型的参数进行估计。本章主要讨论了如何对估计的模型进行统计推断。虽然，双变量模型是最简单地的线性回归模型，但在这两章中介绍的基本思想是后面所要讨论的多元回归模型的基础。我们将会发现，在许多方面，多元回归模型是双变量模型直接的推广。

习题

6.1 解释概念

(a) 最小二乘 (b) OLS估计量 (c) 估计量的方差 (d) 估计量的标准差 (e) 同方差性 (f) 异方差性 (g) 自回归 (h) 总离差平方和(TSS) (i) 回归平方和(ESS) (j) 残差平方和(RSS) (k) 判定系数r^2 (l) 估计值的标准差 (m) BLUE (n) 显著性检验 (o) t检验 (p) 单边检验 (q) 双边检验 (r) 统计显著的

6.2 判断正误并说明理由

(a) OLS就是使误差平方和最小化的估计过程。(b)计算OLS估计量无需古典线性回归模型的基本假定。 (c) 高斯-马尔柯夫定理是OLS的理论依据。 (d) 在双变量回归模型中，若扰动项u_i服从正态分布，则b_2可能是B_2的更为准确的估计值。 (e) 只有当u_i服从正态分布时，OLS估计量b_1,b_2也才服从正态分布。 (f) r^2是TSS/ESS的比值。 (g) 给定显著水平α及自由度，若计算得到的$|t|$值超过临界的t值，我们将接受零假设。 (h) 相关系数r与斜率b_2同号。

6.3 完成下列个空缺

(a) 若$B_2=0$,$b_2/\text{se}(b_2)=\ldots$ (b) 若$B_2=0$,$t=b_2/\ldots$ (c) r^2位于...与...之间。 (e) TSS=RSS+... (f) d.f.(RSS)=d.f (...)+d.f (RSS) (g) $\hat{\sigma}^2$ 称为... (h) $\sum y_i^2 = \sum (Y_i - \ldots)^2$ (i) $\sum \hat{y}_i^2 = b_2(\ldots)$

6.4 考虑下面的回归模型：

$$\hat{Y}_i = -66.105\,8 + 0.065\,0X_i \qquad r^2=0.946\,0$$
$$\text{se}=(10.750\,9) \quad (\qquad) \qquad n=20$$
$$t=(\qquad) \quad (18.73)$$

完成空缺。如果$\alpha=5\%$，你能接受假设：真实的B_2为零吗？你是用单边检验还是双边检验，为什么？

6.5 证明下列r^2的计算公式是恒等的：

$$r^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{b_2^2 \sum x_i^2}{\sum y_i^2} = \frac{\sum (y_i \hat{y}_i)^2}{\left(\sum y_i^2\right)\left(\sum \hat{y}_i^2\right)}$$

6.6 证明：$\sum e_i = n\bar{Y} - nb_1 - nb_2\bar{X} = 0$

6.7 Dale Baies和Larry Peppers[1]根据美国1962~1977年的数据。得到对汽车的需求函数如下：

$$\hat{Y}_t = 5\ 807 + 3.24X_t \qquad r^2 = 0.22$$

$$\text{se} = \qquad (1.63)$$

式中，Y表示零售私人汽车数量(千辆)，X表示真实的可支配收入(以1972年为标准，单位为亿美元。)

注：未给出b_1的标准差，se。

(a) 对B_2建立一个95%的置信区间。 (b)检验假设：该置信区间包括$B_2=0$。如果不包括，你将接受零假设吗？ (c)在$H_0: B_2=0$下，计算t值，在5%的显著水平下，是统计显著的吗？你将选择双边t检验还是单边的？为什么？

6.8 现代投资分析的特征线涉及如下回归方程：

$$r_1 = B_1 + B_2 r_{mt} + u_t$$

其中 r表示股票或债卷的收益率

r_m表示有价证券的收益率(用市场指数表示，比如S&P500)

t表示时间

在投资分析中，B_2称为债券的安全系数β,是用于度量市场的风险程度，即市场的发展对公司财产有怎样的影响。

依据1956～1976年间240个月的数据，Fogler和Ganpathy 得到IBM股票的回归方程；市场指数是用在芝加哥大学建立的市场有价证券指数：[2]

$$\hat{r}_t = 0.726\ 4 + 1.059\ 8 r_{mt} \qquad r^2 = 0.471\ 0$$

$$\text{se} = (0.300\ 1) \quad (0.072\ 8)$$

(a) 解释估计的斜率与截距。 (b) 如何解释r^2？ (c) 安全系数β大于1的证券称为不稳定证券。建立适当的零假设及备择假设，并用t检验进行检验($\alpha=5\%$)。

6.9 下面数据是依据10对X和Y的观察值得到的。

$$\sum Y_i = 1\ 1\ 1\ 0 \ ; \ \sum X_i = 1\ 6\ 8\ 0 \ ; \ \sum X_i Y_i = 204\ 200$$

$$\sum X_i^2 = 315\ 400 \ ; \ \sum Y_i^2 = 133\ 300$$

假定满足所有的古典线性回归模型的假定，求

(a) b_1和b_2？ (b) b_1和b_2的标准差？ (c) r^2？ (d) 对B_1，B_2分别建立95%的置信区间？ (e) 利用置信区间法，你可以接受零假设：$B_2=0$吗？

6.10 依据美国1970～1983年的数据，得到下面的回归结果：

$$\text{GNP}_t = -787.472\ 3 + 8.086\ 3 M_{1t} \qquad r^2 = 0.991\ 2$$

$$\text{se} = (\qquad) \qquad (0.219\ 7)$$

$$t = (-10.0001) \quad (\qquad)$$

其中，GNP是国民生产总值(单位是亿美元)，M_1是货币供给(单位是百万美元)

注：M_1包括现金、活期存款、旅游支票等。

(a) 填充括号内缺省的参数。 (b) 货币学家认为：货币供给对GNP有显著的正面影响，你如何检验这个假设？ (c) 负的截距有什么意义？ (d) 假定1984年m_1为552亿美元，预测该年平均的GNP？

6.11 政治的商业周期：经济事件会影响总统选举吗？为了检验称之为政治商业周期理论，

---

1 参见Dale G.Bails ,Larry C.Peppers，*Business Fluctuations :Forecasting Techniques and Applications*，1982，p.147.

2 参见H.Russell Fogler 与Sundaram Ganapathy，*Financial Econometrics* ，1982，p.13.

Gary Smity[1]依据1928～1980年间，每四年总统选举的数据得到如下回归结果：

$$\hat{Y}_t = 53.10 - 1.70X_t$$
$$t=(34.10) \quad (-2.67) \quad r^2=0.37$$

其中，Y表示收到的公众投票(%)，X表示失业率的变化——选取年的失业率减去上一年的失业率。

(a) 先验地，你预期X的符号为正还是负？ (b) 该回归结果证实了政治商业周期理论吗？写出你的求证过程。 (c) 1984～1988年的总统选取的结果是否验证了该理论？ (d) 如何计算b_1和b_2的标准差？

6.12 为了研究美国制造业设备利用与通货膨胀之间的关系，Thomas A. Gittings[2]根据1971～1988每年的数据得到下面的回归结果：

$$\hat{Y}_t = -70.85 + 0.888\,0X_t$$
$$t=(-5.89) \quad (5.90) \quad r^2=0.685$$

其中，Y表示通货膨胀的变化(以批发价格指数来度量)，X表示设备利用率。

(a) 先验地，你认为X与Y正相关吗？为什么？ (b) 如何解释该回归方程中的斜率？ (c) 估计的斜率值是统计上显著的吗？ (d) 它是否显著不为1？ (e) 定义设备自然利用率为当Y为零时的X的值。求设备自然利用率。

6.13 表6-6给出了1974~1986年间制造业中的税后利润X(亿美元)以及三月期现金利息Y(亿美元)的数据。

表6-6 美国制造业中现金利息(Y)与税后利润(X)，

年份	Y	X	年	Y	X
	百万美元			百万美元	
1974	19 467	58 747	1981	40 317	101 302
1975	19 968	49 135	1982	41 259	71 028
1976	22 763	64 519	1983	41 624	85 834
1977	26 585	70 366	1984	45 102	107 648
1978	28 932	81 148	1985	45 517	87 648
1979	32 491	98 698	1986	46 044	83 121
1980	36 495	92 579			

资料来源：Business Statistics, 1986, U.S.Department of Commerce, Bureau of Economic Analysis, December 1987, p.72.

(a) 你预期现金与税后利润的关系如何？ (b) 作Y与X间的散点图 (c) 该散点图是否与你的预期相符？ (d) 如果是的，运用OLS进行回归，并求出一般统计量。 (e) 对真实斜率建立一个99%置信区间并检验假设：真实的斜率为零；即现金利息与税后利润之间不相关。

表6-7 真实小时工资指数与小时产出指数 (1997=100)

年份	Y	X	年	Y	X
1973	96.8	95.9	1981	95.8	100.7
1974	95.4	93.9	1982	97.3	100.3
1975	96.0	95.7	1983	98.2	103.0
1976	98.8	98.3	1984	97.9	105.5
1977	100.0	100.0	1985	98.8	107.7
1978	100.9	100.8	1986	101.2	110.1
1979	99.4	99.6	1987	101.5	111.0
1980	96.7	99.3			

资料来源：Economic Report of the President,1989,Table B-46,p.360.

1 Gary Smity，《统计推论》(Statistical Reasoning)，1985，第488页。符号略有改动。原始数据出自，Ray C.Fair“经济事件对总统选举的影响”，《经济与统计评论》(*The Review of Economics and Statistics*)，1978，5，第159~173页。

2 Thomas A. Gittings, “Capacity Utilization and Inflation,” *Economic Perspectives*, Federal Reserve Bank of Chicago, May/June 1989,p.4.

6.14 **收入－生产率关系。**

表6-7**给出**1973~1987**年间美国商业部提供的小时真实工资数**Y(1997=100)**与小时产出**X**。**

(a) **经济理论表明收入与生产率之间存在正相关。作出** X**与**Y**间的散点图看是否支持该理论？** (b) **如果是，运用普通最小二乘法进行回归分析。** (c) **是收入是生产力的函数还是生产力是收入的函数？你是如何决定的？**

表6-8 **证券收益率**(Y)**与市场收益率**(X)

Y	X	Y	X
67.5	19.5	19.2	8.5
−35.2	−29.3	−42.0	−26.5
63.7	61.9	19.3	45.5
3.6	9.5	20.0	14.0
40.3	35.3		

6.15 **通过原点的回归：无截距的回归模型。有下列一种情况，假定的变量之间存在如下形式的总体回归函数：**

$$Y_i = B_2 X_i + u_i$$

在这个模型中,截距为零。因此,称此模型为通过原点的回归模型或零一截距模型。可以证明：

$$b_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$\text{var}(b_2) = \frac{\sigma^2}{\sum X_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-1}$$

注意：由于回归经过原点，因此,这里不能再用常用的判定系数 r^2**的计算公式。**[1]

(a) **与双变量模型相比较，指出它们的区别？** (b) **为什么在此模型中的自由度是**(n－1)**而不是**(n－2)? (c) **考虑表**6-8**中的数据,其中**Y**表示证券的收益率，**X**表示市场指数的收益率**(10**年为一期)。**

(1) **对这些数据拟合出一条经过原点的直线。**

(2) **拟合出有截距的回归模型。**

(3) **根据上述结果，你认为哪个模型更好？为什么？**

6.16 **式**(6-52)**的进口支出与**PDI**回归的原始数据如下：利用这些数据，验证式**(6-32)**给出的回归结果。由于舍入误差，计算出的两个结果会稍有不同。**

$$\bar{Y} = 253.08 \qquad \bar{X} = 2\,096.7 \qquad \sum (Y_i - \bar{Y})^2 = 139.830$$

$$\sum (X_i - \bar{X})^2 = 2\,182\,900 \qquad \sum (Y_i - \bar{Y})(X_i - \bar{X}) = 535\,300\text{；} \qquad n = 20$$

6.17 **参考表**5-5**中给出的**S.A.T**数据。假定想根据下面的回归方程，通过女生的数学分数来预测男生的数学分数：**

$$Y_t = B_1 + B_2 X_t + u_t$$

其中，Y**，**X**分别代表男，女生的数学分数。**

(a) **估计上述回归方程，求常用的统计量。** (b) **检验假设：**Y**与**X**不相关。** (c) **假定**1991**年，预期女生的数学分数为**460**，预测该年男生的平均分数。** (d) **对**(c)**的预测值建立一**

1 **参见**Dennis J.Aigner,*Basic Econometrics*，1971，pp.85-90。

个95%的置信区间。

6.18 重复习题6.17，但令Y和X分别代表男、女生词汇分数。这里假定1991年女生的词汇分数为425。

6.19 考虑下面的回归结果：[1]

$$\hat{Y}_t = -0.17 + 5.26X_t \qquad \overline{R}^2 = 0.10 \qquad \text{D-W} = 2.01$$

其中，Y表示本年度1月份到次年1月份股票价格指数的实际收益

X表示上一年总股息与上一年股票价格指数之比

t表示时间

注：1. D-W统计量，参见第12章。

时间为1926~1982年。

2. $\overline{R}^2$表示经过校正的判定系数。D-W值是用以度量自相关的指标，我们将在随后的章节中作出解释。

(a)如何解释上述回归方程？ (b)如果你接受上面的结论，是不是意味着：当股息/价格的比率高时，最好的投资战略是对股票进行投资？ (c)如果你想知道(b)的答案，参阅Shille's的分析。[2]

1 参见Robert J.Shiller,*Market Volatity*，MIT Press，1989，pp.32-36。

2 Ibid .

第 7 章

多元回归：估计与假设检验

在双变量线性回归模型中，我们仅考虑了一个自变量或解释变量。本章我们将把模型扩展到有多个解释变量影响应变量的情形。包含多个解释变量的回归模型，称为多元回归模型(Multiple Regression Model)。多元是指有多种因素(即变量)对应变量有影响。

举个例子，考虑80年代在某些州出现的由于存贷款机构的破产所导致的存贷危机。假定我们想要建立一个模型来解释破产这个应变量。像破产这样的现象很复杂以至很难仅仅只用一个变量来解释它，因此需要多个解释变量。比如，原始资本与总资产的比率，超过 90天的贷款与总资产的比率，非自然增加贷款与总资产的比率，重议价贷款与总资产的比率；净收入与总资产的比率等等。[1]

勿庸赘言，我们可以举出许多多元回归的例子。实际上，许多回归模型都是多元回归模型，因为很少有经济现象能够仅用一个解释变量就能解释清楚。

本章讨论多元回归模型旨在探求下列问题的答案：

(1) 如何估计多元回归模型？多元回归模型的估计过程与双变量模型有何不同？

(2) 对多元回归模型的假设过程与双变量模型有何不同？

(3) 多元回归有没有一些在双变量模型中未曾遇到过的独特的特性？

(4) 既然一个多元回归模型能够包括任意多个解释变量，那么，对于具体的情况，我们如何决定解释变量的个数？

为了回答这些以及与此相关的问题，我们先考虑最简单的多元回归模型，三变量模型，也即在此模型中，有一个应变量 Y，两个自变量 X_2、X_3。一旦清楚地理解了三变量模型，就很容易将模型扩展到四个、五个或更多个变量的情形，只不过计算复杂了一些而已(但在计算机时代，这并不是一个什么可怕的问题)。有意思的是，三变量模型本身在许多方面就是双变量模型的直接扩展。从下面的讨论中不难看到这一点。

7.1 三变量线性回归模型

将双变量总体回归模型(PRF)推广，便可写出不含随机项的三变量总体回归模型：

$$E(Y_t)=B_1+B_2X_{2t}+B_3X_{3t} \qquad (7\text{-}1)^2$$

1 事实上，这是联邦储备在对破产银行的研究中考虑的几个变量。

2 方程(7-1)还可写为：$E(Y_t)=B_1X_{1t}+B_2X_{2t}+B_3X_{3t}$，这里可理解为对每个观察值，$X_{1t}=1$。方程(7-1)中为了符号的方便，参数的下标与估计量的下标相一致。

其随机形式为：

$$Y_t=B_1+B_2X_{2t}+B_3X_{3t}+u_t \tag{7-2}$$

$$=E(Y_t)+u_t \tag{7-3}$$

式中Y—— 应变量；X_2、X_3—— 解释变量；u—— 随机扰动项；t—— 第t个观察值。

在此例中的数据是横截面数据，下标i表示了第i个观察值。注意，我们将随机项u引入到三变量模型中，或者更一般地，引入到多元回归模型中，其原因与第 2章讨论的双变量的情形相同。

B_1是截距。与以前一样，它表示了当X_2、X_3为零时的Y平均值。B_2、B_3称为偏回归系数，随后解释它们的含义。

根据第5章的讨论，我们知道式(7-1)给出了在给定X_2、X_3值的情况下的Y的条件均值。因此，与双变量情形相同，多元回归分析也是条件回归分析，是在给定解释变量的值的条件下，得到Y的均值：回顾总体回归函数(PRF)给出了在给定解释变量X_2、X_3时的Y总体的(条件)均值。[1]

多元模型随机的形式，式(7-2)，表明任何一个Y值可以表示成为两部分之和：

(1) 系统成分或决定成分，$(B_1+B_2X_{2t}+B_3X_{3t})$，也就是$Y$的均值$E(Y_t)$，(即回归线上的点)。[2]

(2) 非系统成分u_t，是由除X_2、X_3以外其他因素决定的。

所有这些都与双变量情形类似；惟一需要强调的是：现在的解释变量有两个而不是一个。

注意：模型(7-1)或相对应的随机形式(7-2)是一个线性回归模型 —— 即模型的参数B之间是线性的。在第5章讲过，本书仅仅关注参数之间是线性的情况；在这类线性模型中，变量之间可能是线性的，也可能不是(更详细的讨论见第８章)。

偏回归系数的含义

前面讲到，B_1,B_2称为偏回归系数(partial regression coefficients)或偏斜率系数(partial slope coefficients)。其意义如下：B_2度量了在X_3保持不变的情况下，X_2每变动一单位，Y的均值$E(Y)$的改变量。同样的，B_3度量了在X_2保持不变的情况下，X_3每变动一单位，Y的均值$E(Y)$的改变量。这是多元回归一个特殊的性质；在双变量情形下，由于仅有一个解释变量，无须担心模型中出现其他变量。而在多元回归中，我们想要知道Y平均值的变动有多大比例“直接”来源于X_2的变动，多大比例“直接”来源于X_3的变动。这一点对于理解多元回归的内在逻辑很重要。我们以一个简单的例子来说明，假定有如下总体回归函数：

$$E(Y_t)=15-1.2X_{2t}+0.8X_{3t} \tag{7-4}$$

令X_3取值为10，将其代入式(7-4)，得

$$\begin{aligned}E(Y_t)&=15-1.2X_{2t}+0.8(10)\\&=(15+8)-1.2X_{2t}\\&=23-1.2X_{2t}\end{aligned} \tag{7-5}$$

这里，斜率$B_2=-1.2$表示当X_3为常数时，X_2每增加一个单位，Y的平均值将减少1.2个单位 —— 在这个例子中，X_3为常数10(若取其他常数也一样)。[3]这个斜率就称为偏回归系数。[4]同样地，如果X_2为常数，比如说，$X_2=5$，得到

$$\begin{aligned}E(Y_t)&=15-1.2(5)+0.8X_{3t}\\&=9+0.8X_{3t}\end{aligned} \tag{7-6}$$

---

1 与双变量模型不同的是，我们不能给出其图形，因为三个变量Y、X_2、X_3，需用三维图，很难设想在二维平面上画出三维图会是怎样的。但展开联想，可以勾画出形如 5-1的图形。

2 从图形上看，这种情形下的总体回归线表示了一个平面。

3 可以证明，无论X_3取任何常数，结果都一样。因为常数与其系数的乘积仍为常数，只不过增加了截距值。

4 数学基础较好的学生，立刻可以看出B_2是$E(Y)$关于X_2的偏导数，B_3是$E(Y)$关于X_3的偏导数。

这里，斜率$B_3=0.8$，表示当X_2为常量时，X_3每增加1个单位，Y的平均值将增加0.8个单位，如果X_2为其他常量，结果不变。这个斜率系数也是偏回归系数。

简言之，偏回归系数反映了当模型中的其中一个解释变量为常量时，另一个解释变量对应变量均值的影响。多元回归的这个独特性质不但能使我们引入多个解释变量，而且能够“分离”出每个解释变量X对应变量Y的影响。

7.2 多元线性回归模型的若干假定

与双变量模型相同，首先要对多元回归模型的参数进行估计。为了这个目的，我们仍在第6章介绍的古典线性回归模型的框架下利用普通最小二乘法(OLS)对参数进行估计。

特别地，对模型(7-2)，作如下假定：(参见6.1节)

A7.1 X_2、X_3与扰动项u不相关。但是，如果X_2、X_3是非随机变量(也即X_2、X_3在重复抽样中取某固定数值)，这个假定将自动满足。

由于我们的回归分析是条件回归分析，即在给定X值条件下的回归分析，因此假定7.1是不必要的。但是在处理第15章中的联立方程回归模型时，将会看到有些变量可能与误差项相关。

A7.2 零均值假定：

$$E(u_i)=0 \tag{7-7}$$

A7.3 同方差假定，即u的方差为一常量：

$$\mathrm{Var}(u_i)=\sigma^2 \tag{7-8}$$

B7.4 无自相关假定

$$\mathrm{cov}(u_i,u_j)=0,i\neq j \tag{7-9}$$

A7.5 解释变量之间不存在线性相关关系。即两个解释变量之间无确切的线性关系，这是一个新的假定，我们随后解释。

A7.6 为了假设检验，假定随项误差u服从均值为零，(同)方差为σ^2的正态分布。即，

$$u_i\sim N(0,\sigma^2) \tag{7-10}$$

除了假定7.5外，其他的假定的基本原理都与前面讨论的双变量线性回归模型相同，正如在前面章节中提到的那样，这些假定是为了保证能够使用OLS法来估计模型的参数。在第三部分中，我们将重新检查这些假定，看看若在实际中其中一条或几条假定不满足时，会发生什么情况。

假定7.5表明了解释变量X_2与X_3之间不存在完全的线性关系，用统计学语言，称为非共线性(no collinearity)或非多重共线性(no multicolliearity)。

一般地，非完全共线是指变量X_2不能表示为另一变量X_3的完全线性函数。因而，若有：

$$X_{2i}=3+2X_{3i}$$

或

$$X_{2i}=4X_{3i}$$

则这两个变量之间是共线性的，因为X_2、X_3之间存在完全的线性关系，假定7.5表明不存在共线性。逻辑很简单，举个例子，如果$X_2=4X_3$，则将它代入式(7-1)，会发现：

$$\begin{aligned}E(Y_i)&=B_1+B_2(4X_{3i})+B_3X_{3i}\\&=B_1+(4B_2+B_3)X_{3i}\\&=B_1+AX_{3i}\end{aligned} \tag{7-11}$$

式中，

$$A=4B_2+B_3 \tag{7-12}$$

模型(7-11)是一个双变量模型，而非三变量模型。即使能够对模型(7-11)估计，也无法从估计的A值得到B_2、B_3的值。注意：方程(7-12)是有两个未知数的方程，而求B_2和B_3的估计值需要两个(独立)的方程。

我们得出的结论是：在存在完全共线性的情况下，不能估计偏回归系数B_2和B_3的值；换句话说，不能估计解释变量X_2和X_3各自对因变量Y的影响。但是，这也没有什么奇怪的，因为在模型中确实没有两个独立的解释变量。

虽然在实际中，很少有完全共线性的情况，但是高度完全共线性(high perfect collinearity)或近似完全共线性(near perfect collinearity)的情况还是很多的。在后面的章节中(参见第10章)我们将详细讨论这类情况。现在仅考虑两个或多个解释变量之间不存在完全的线性关系的模型。

7.3 多元回归参数的估计

我们用OLS法估计模型(7-2)的参数，有关OLS法的一些性质已在第5章和第6章中讨论过。

7.3.1 普通最小二乘估计量

要求OLS估计量，首先需写出与总体回归模型(7-2)相对应的样本回归模型，

$$Y_t=b_1+b_2X_{2t}+b_3X_{3t}+e_t \tag{7-13}$$

其中，e为残差项，简称残差(与总体回归模型中的误差项u相对应)，b是总体系数B的估计量。更具体地，

b_1＝B_1的估计量　　b_2=B_2的估计量　　b_3=B_3的估计量

样本回归方程：

$$\hat{Y}_t=b_1+b_2X_{2t}+b_3X_{3t} \tag{7-14}$$

方程(7-14)是估计的总体回归线(实际上是一个平面)。

在第5章中已经解释过，OLS原则是选择未知参数值使得残差平方和(RSS)$\sum e_i^2$尽可能小。首先，将模型(7-13)重写：

$$e_t=Y_t-b_1-b_2X_{2t}-b_3X_{3t} \tag{7-15}$$

同时将两边平方再求和，得

$$\text{RSS}：\sum e_t^2=\sum\left(Y_t-b_1-b_2X_{2t}-b_3X_{3t}\right)^2 \tag{7-16}$$

最小二乘法就是使RSS(Y_i的真实值与估计值之差平方和)最小化。

式(7-16)最小化过程将用到偏微分。这里不再详细推导，得到下面的正规方程：[1][比较与之相对应的双变量模型的正规方程(5-14)和(5-15)。]

$$\bar{Y}=b_1+b_2\bar{X}_2+b_3\bar{X}_3 \tag{7-17}$$

$$\sum Y_tX_{2t}=b_1\sum X_{2t}+b_2\sum X_{2t}^2+b_3\sum X_{2t}X_{3t} \tag{7-18}$$

其中，求和符号表示是从1到n。这里有三个方程，三个未知数——已知应变量Y和两个解释变量X的值，b未知。通常可由三个方程求解三个未知数。对上面方程作简单的代数变换，得到三个OLS估计量的表达式如下：

1 详细的数学推导参见附录7A。

$$b_1 = \bar{Y} - b_2\bar{X}_2 - b_3\bar{X}_3 \tag{7-20}$$

$$b_2 = \frac{\left(\sum y_t x_{2t}\right)\left(\sum x_{3t}^2\right) - \left(\sum y_t x_{3t}\right)\left(\sum x_{2t}x_{3t}\right)}{\left(\sum x_{2t}^2\right)\left(\sum x_{3t}^2\right) - \left(\sum X_{2t}x_{3t}\right)^2} \tag{7-21}$$

$$b_3 = \frac{\left(\sum y_t x_{3t}\right)\left(\sum x_{2t}^2\right) - \left(\sum y_t x_{2t}\right)\left(\sum x_{2t}x_{3t}\right)}{\left(\sum x_{2t}^2\right)\left(\sum x_{3t}^2\right) - \left(\sum x_{2t}x_{3t}\right)^2} \tag{7-22}$$

其中，小写字母表示其值与其样本均值之差。(例如，$y_t = Y_t - \bar{Y}$)

读者可能会注意到上述表达式与双变量模型中给出的相应的表达式(5-16)(5-17)相类似。同样的，这些方程有如下特征：(1)式(7-21)和式(7-22)是对称的，将式(7-22)中的X_2换为X_3即得式(7-22)。(2)式(7-21)与式(7-22)的分母相同。

7.3.2 OLS估计量的方差与标准差

得到截距及偏回归系数的OLS估计量之后，我们就可以推导出这些估计量方差及标准差(与双变量模型相同)。这些方差或标准差表示了估计量由于样本的改变而发生的变化。与双变量模型相同，需要标准差是出于两个主要目的：(1)建立真实参数值的置信区间。(2)检验统计假设。下面给出相关的公式，证明略去：

$$\text{var}(b_1) = \left[\frac{1}{n} + \frac{\bar{X}_2^2\sum x_{3t}^2 + \bar{X}_3^2\sum x_{2t}^2 - 2\bar{X}_2\bar{X}_3\sum x_{2t}x_{3t}}{\sum x_{2t}^2\sum x_{3t}^2 - \left(\sum x_{2t}x_{3t}\right)^2}\right]\sigma^2 \tag{7-23}$$

$$\text{se}(b_1) = \sqrt{\text{var}(b_1)} \tag{7-24}$$

$$\text{var}(b_2) = \frac{\sum x_{3t}^2}{\left(\sum x_{2t}^2\right)\left(\sum x_{3t}^2\right) - \left(\sum x_{2t}x_{3t}\right)^2}\sigma^2 \tag{7-25}$$

$$\text{se}(b_2) = \sqrt{\text{var}(b_2)} \tag{7-26}$$

$$\text{var}(b_2) = \frac{\sum x_{2t}^2}{\left(\sum x_{2t}^2\right)\left(\sum x_{3t}^2\right) - \left(\sum x_{2t}x_{3t}\right)^2}\sigma^2 \tag{7-27}$$

$$\text{se}(b_3) = \sqrt{\text{var}(b_3)} \tag{7-28}$$

在所有这些表达式中，σ^2表示总体误差项u_t的(同)方差，其OLS估计量：

$$\hat{\sigma}^2 = \frac{\sum e_t^2}{n-3} \tag{7-29}$$

式(7-29)是双变量模型[式(6-8)]的直接的扩展，只不过此时的自由度为$n-3$。这是因为在估计RSS，$\sum e_t^2$时，必须先求出b_1、b_2、b_3，也就是说，它们“消耗”了三个自由度。以此类推，在4个解释变量情形下，自由度为($n-4$)；当解释变量是5个时，自由度为($n-5$)；需要提醒注意的是b_2的正的平方根是估计值的标准差或称回归标准差(即Y值偏离估计回归线的标准差)，即：

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} \tag{7-30}$$

如何计算$\sum e_t^2$呢？由于$\sum e_t^2 = \left(Y_t - \hat{Y}_t\right)^2$，因此在计算$\sum e_t^2$时，首先要求$\hat{Y}_t$，我们很容易求出$\hat{Y}_t$。但计算RSS有一个更简便的方法(见附录7A.2)，

$$\sum e_t^2 = \sum y_t^2 - b_2 \sum y_t x_{2t} - b_3 \sum y_t x_{3t} \tag{7-31}$$

也就是说，一旦估计出偏斜率的值，就很容易求得 $\sum e_t^2$ 。

7.3.3 多元回归OLS估计量的性质

在双变量模型中，我们看到在古典线性回归模型的基本假定下， OLS估计量是最优线性无偏估计量。这个性质对于多元回归同样成立。因此，根据 OLS估计的每一个回归系数都是线性的和无偏的 —— 平均而言，它与真实值相一致。在所有线性无偏估计量中， OLS估计量具有最小方差性，所以OLS估计量比其他线性无偏估计量更准确地估计了真实的参数值。

从上面的讨论中，不难发现，三变量模型在许多方面是双变量模型的推广，只不过估计公式略显复杂。解释变量的个数如果多于三个，那么得到的计算公式将会更复杂。在那种情况下，不管你愿不愿意，都必须用矩阵代数来计算。当然，在本书中将不会遇到用矩阵代数计算的情况，不过现在很少有人徒手计算，还是让计算机做这些复杂的工作吧！

7.4 实例：未偿付抵押贷款债务 (美国：1980～1985年)

我们花一些时间用一个数字例子来说明上面讨论的理论。看表 7-1提供的数据。这些数据包括非农业未尝付抵押贷款(Y，亿美元)，个人收入(X_2，亿美元)，新住宅抵押贷款费用(X_3，%)，抵押贷款费用包括对常规抵押贷款和手续费收取的利率。为了讨论的方便，我们把这些变量简称为抵押债务(Y)，收入(X_2)，抵押费用(X_3)。

先验地，预期抵押债务与收入正相关，因为个人收入越高，则其借贷购买新房的能力就越强，因此，X_2的系数B_2预期为正；预期抵押债务与抵押费用负相关，原因很简单，在其他条件保持不变时，如果购房费用上升，则对住房的需求将下降，从而减少了对新的抵押贷款的需求。相应地，X_3的系数B_3预期为负。我们将会看到，这些先验的或理论的预期通常会有助于对回归分析的结果进行估计。

表7-1 美国抵押贷款债务，个人收入，抵押贷款费用

年 份	非农业抵押贷款债务/亿美元	个 人 收 入/亿美元	新抵押贷款费用(%)
1980	1365.5	2285.7	12.66
1981	1465.5	2560.4	14.70
1982	1539.3	2718.7	15.14
1983	1728.2	2891.7	12.57
1984	1958.7	3205.5	12.38
1985	2228.3	3439.6	11.55
1986	2539.9	3647.5	10.17
1987	2897.6	3877.3	9.31
1988	3197.3	4172.8	9.19
1989	3501.7	4489.3	10.13
1990	3723.4	4791.6	10.05
1991	3880.9	4968.5	9.32
1992	4011.1	5264.2	8.24
1993	4185.7	5480.3	7.20
1994	4389.7	5753.1	7.49
1995	4622.0	6115.1	7.87

资料来源：《总统经济报告》(Economic Report of the Present)，1997年，抵押贷款债务的数据摘自表B-73，第386页，个人收入的数据摘自表B-28，第332页，新抵押贷款费用的数据摘自B-71，第382页。

定义了Y，X_2，X_3，现在想要用OLS法估计式(7-2)。首先需要知道这些变量的各种平方和与交差乘积和的值。表 7-2给出了这些必要的数据。当然，现在已不用徒手计算了。在附录7A.4中，我们给出根据表7-1的数据得到的回归结果[模型(7-2)]。这里给出用EVIEWS统计软件得到的回归结果，我们将逐步地对回归结果作出解释。

表7-2 回归方程(7-32)的原始数据

$\sum Y_t = 47\ 235.2$	$\sum X_{2t} = 65\ 660.8$	$\sum X_{3t} = 167.968$
$\sum y_t^2 = 19\ 223\ 091$	$\sum x_{2t}^2 = 21965732.96$	$\sum x_{3t}^2 = 1742.89$
$\sum y_t x_{2t} = 20\ 411719$	$\sum y_t x_{3t} = -38.337$	$\sum x_{2t} x_{3t} = -40293$
$n=16$		

7.4.1 回归结果

根据附录7A.4给出的计算结果，得到下列回归方程。(注：这里保留小数点后四位数)首先用式(6-48)的形式写出回归结果，然后按步骤讨论：

$$\hat{Y}_t = 155.681\ 2 + 0.825\ 8X_{2t} - 56.439\ 3X_{3t} \qquad (7\text{-}32)$$

$$\begin{aligned} \text{se} &= (578.328\ 8) \quad (0.063\ 5) \quad (31.454\ 3) \\ t &= (0.269\ 2) \quad (12.991\ 0) \quad (-1.794\ 3) \qquad R^2 = 0.989\ 4 \\ p\text{值} &= (0.792\ 0) \quad (0.000\ 0)^* \quad (0.096\ 0) \text{校正的} \ R^2 = 0.987\ 8 \end{aligned}$$

*表示很小，几乎为零。

在计算t值时，假定相应的总体系数为零。

7.4.2 对回归结果的解释

回归结果表明：X_2的偏回归系数0.824 8表示了在其他变量(也即抵押贷款费用)保持不变时，收入每增加1美元，抵押贷款债务平均地将增加38美分。与预期相同，这两个变量之间正相关。

同样地，部分斜率系数－56.439 4表示在其他变量(也即收入)保持不变时，抵押贷款费用每上升1个百分点，则平均的抵押贷款债务将下降56亿美元。与预期相同，这两个变量之间负相关。

截距155.681 2表示如果收入和抵押贷款费用都为零时，平均的抵押贷款债务量约为 157亿美元。但是，在本例中，截距没有什么经济意义。

对估计的抵押贷款债务函数，可以得到什么结论呢？到目前这个阶段，我们仅能够说估计的函数有经济意义。但是，估计参数的统计显著性如何呢？估计的回归方程的优度如何呢？我们将在下面的分析中给出答案。

7.5 估计的多元回归方程的拟合优度：多元判定系数 R^2

在双变量模型中，我们知道式(6-40)定义的r^2是用来度量拟合的样本回归直线(SRL)的拟合优度；也就是说，r^2给出了单个解释变量 X对应变量 Y变动的解释比例或解释的百分比。 r^2的概念可以推广到包含若干个解释变量的回归模型之中。因此，在三变量模型中，我们想知道， X_2和X_3一起对应变量 Y变动的解释程度(解释比例)。我们将度量这个信息的量称为多元判定系数(Multiple Coefficient of Determination)，用符号R^2表示；从概念上讲，它与r^2类似。

与双变量模型相同，有如下恒等式：

$$\text{TSS} = \text{ESS} + \text{RSS} \qquad (7\text{-}33)$$

其中，

TSS=总离差平方和$(=\sum y_t^2)$

ESS＝回归平方和

RSS＝残差平方和

同样地，与双变量模型类似，R^2定义如下：

$$R^2 = \frac{\text{ESS}}{\text{TSS}} \tag{7-34}$$

即，R^2是回归平方和与总离差平方和的比值；与双变量模型惟一不同的是现在的 ESS值与多个解释变量有关。

可以证明：[1]

$$\text{ESS} = b_2 \sum y_t x_{2t} + b_3 \sum y_t x_{3t} \tag{7-35}$$

前面已经给出：

$$\text{RSS} = \sum y_t^2 - b_2 \sum y_t x_{2t} - b_{3t} \sum y_t x_{3t} \tag{7-31}$$

因此，

$$R^2 = \frac{b_2 \sum y_t x_{2t} + b_3 \sum y_t x_{3t}}{\sum y_t^2} \tag{7-36}$$[2]

所有这些量都可以根据表 7-2提供的数据求得。当然，计算机将会很快地求得结果。根据附录7A.4的计算结果，R^2为0.989 4，它表示约99%的抵押贷款债务的变化可以由个人收入和抵押贷款费用这两个变量来解释。与 r^2相同，R^2的值也在0与1之间(为什么？)R^2越接近于1，表示估计的回归直线拟合得越好；R^2近似等于1，表示回归直线非常好地拟合了样本数据。

R^2的正的平方根，称为多元相关系数(coefficient of multiple correlation)，与双变量模型的r相类似。正如r度量了Y与X的线性相关程度一样，R度量了Y与所有解释变量的线性相关程度。虽然r可正可负，但R却总取正值。但是，在实际中，很少用到R。在本例中，$R = \sqrt{0.9894} = 0.9947$。

7.6 多元回归的假设检验：一般的解释

虽然R^2度量了估计的回归直线的拟合优度，但是R^2本身却不能告诉我们估计的回归系数是否在统计上是显著的，即是否显著不为零。有的回归系数可能是显著的，有些则可能不是。如何判断呢？

具体地，假定想要检验假设：个人收入对抵押贷款债务没有影响。换句话说,要检验零假设：H_0:B_2=0。如何进行该假设检验呢？根据第6章中讨论的对双变量模型的假设检验,要回答这个问题，需要求B_2的估计量b_2的抽样分布。那么b_2的抽样分布是什么呢？b_1和b_2的抽样分布又是什么呢？

在双变量模型中，如果假定误差项 u服从正态分布，则OLS估计量b_1、b_2服从正态分布。在A7.6假定中我们已经规定了即使对多元回归，仍假定 u服从均值为0，方差为σ^2的正态分布。在这个假定以及7.2节列出的其他基本假定满足的条件下，可以证明，b_1、b_2、b_3均服从均值分别为B_1、B_2、B_3正态分布。式(7-23)，(7-25)和(7-27)分别给出了它们的方差。

但是，与双变量模型相同，如果用真实的但不可观察的 σ^2的无偏估计量 $\hat{\sigma}^2$代替σ^2，则OLS估计量服从自由度为$(n-3)$的t分布，而不是正态分布。即

---

1 参见附录7A.2。

2 R^2也可以公式：$1 - \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{\sum e_t^2}{\sum y_t^2}$ 来计算。

$$t = \frac{b_1 - B_1}{\mathrm{se}(b_1)} \sim t(n-3) \tag{7-37}$$

$$t = \frac{b_2 - B_2}{\mathrm{se}(b_2)} \sim t(n-3) \tag{7-38}$$

$$t = \frac{b_3 - B_3}{\mathrm{se}(b_3)} \sim t(n-3) \tag{7-39}$$

注意：此时的自由度为$(n-3)$，因为在计算RSS，$\sum e_i^2$，即而$\hat{\sigma}^2$时，首先需要估计截距及两个部分斜率，也就是说，失去了3个自由度。

用$\hat{\sigma}^2$代替σ^2使得OLS估计量服从t分布，同样在建立置信区间以及对真实的部分斜率进行假设检验时，也可以用$\hat{\sigma}^2$代替σ^2。其内在机制在许多方面与双变量模型相类似，我们用抵押贷款债务一例来说明。

7.7 对回归参数进行假设检验

假定对式(7-32)的回归结果，做如下假设：

$$H_0：B_2=0，H_1：B_2\neq 0$$

也即，在零假设下，个人收入对贷款债务无影响。在备择假设下，个人收入对抵押贷款债务有(正的或负的)影响。因此备择假设是双边假设。

在上述零假设下，我们知道：

$$t = \frac{b_2 - B_2}{\mathrm{se}(b_2)} \tag{7-38}$$

$$= \frac{b_2}{\mathrm{se}(b_2)} \quad (\text{注：} B_2=0)$$

即服从自由度为$(n-3)=13$的t分布。根据附录7A.4中的数据，可得

$$t = \frac{0.82\ 582}{0.0635\ 7}$$

$$=12.9\ 910 \tag{7-40}$$

服从自由度为13的t分布。

根据计算的t值，能否拒绝零假设：个人收入对抵押贷款债务没有影响。要回答这个问题，可用置信区间法或显著性检验法(与双变量模型讨论的相同)。

7.7.1 显著性检验法

回忆一下，在显著性检验方法中，我们需要建立一个统计量，求其抽样分布，选择一个显著水平α，并决定在所选显著水平下检验统计量的临界值。然后将从样本得到的统计量值与其临界值作比较，如果统计量的值超过临界值，则拒绝零假设。[1] 我们可以将这种检验方法推广到多元回归模型中。

回到上面的例子，我们知道在这个问题中，检验统计量是 t统计量，它服从自由度为$(n-3)$的t分布，因此，我们用t显著性检验。实际的机制很简单。假定选择$\alpha=0.05$或5%。由于备择假设是双边的，因此求得的t临界值是在$\alpha/2=2.5\%$的显著水平下(为什么？)。此时的自由度为13。查t分布表得：

$$P(-2.160\leq t\leq 2.60)=0.95 \tag{7-41}$$

1 如果检验统计量是一个负值，那么就考虑其绝对值，如果其绝对值超过临界值，则拒绝零假设。

即t值位于临界值－2.160与2.160之间的概率为95%。

根据式(7-40)可知，在零假设H_0:B_2=0下，计算的t值接近13，显然超过t临界值2.160。因此，拒绝零假设并得出结论：个人收水对抵押贷款债务有影响。

1. p值

实际上，我们无需选择显著水平 α，因为在附录7A.4的结果中已经给出估计的 t统计量的p值。在本例中，p值为0.000 0，几乎为零。因此，如果零假设 B_2＝0为真，那么得此 t值大于或等于13的概率几乎为零。这就是为什么能够拒绝零假设的原因。事实上，根据p值比根据任选的显著水平α更有理由拒绝零假设。[1]

2. 单边或双边检验

由于先验地，预期收入系数为正，因此，这里实际上用得是单边检验。在 5%的显著水平下，单边检验的 t值为1.771。因为计算的 t值约为13仍远大于1.771，所以拒绝零假设并得出结论：个人收入对抵押贷款债务有正向冲击。另一方面，双边检验的结果只是简单地告诉我们个人收入对抵押贷款债务可能有正向的或反向的影响。因此，一定要注意你是如何规定零假设和备择假设的：在选择这些假设时要以经济理论为依据。

7.7.2 置信区间法

假设检验的置信区间法的基本思想已在前面的章节中讨论过。这里，我们仅以一个数字例子来说明。前面已经得到：

$$P(-2.160 \leq t \leq 2.160)=0.95 \tag{7-41}$$

还知道：

$$t=\frac{b_2-B_2}{\mathrm{se}(b_2)} \tag{7-38}$$

如果将(7-38)式代入到(7-41)式，得，

$$p\left\{-2.160 \leqslant \frac{b_2-B_2}{\mathrm{se}(b_2)} \leqslant 2.160\right\}=0.95$$

整理得：

$$p\left[b_2-2.160\mathrm{se}(b_2) \leqslant B_2 \leqslant b_2+2.160\mathrm{se}(b_2)\right]=0.95 \tag{7-42}$$

这就是在7.5%显著水平下B_2的置信区间。回忆一下置信区间法，如果置信区间(也称接受区域)包括了零假设值，那么将不能拒绝零假设。另一方面，如果零假设值在置信区间之外，则拒绝零假设。但需时刻注意的是，无论做何种决定，你犯错误的概率为 5%。

对本例，相应的式(7-42)为

$$0.825\ 8-2.160(0.063\ 6) \leq B_2 \leq 0.825\ 8+2.160(0.063\ 6)$$

即，

$$0.688\ 5 \leq B_2 \leq 0.963\ 1 \tag{7-43}$$

即为B_2的95%的置信区间。由于该置信区间不包括零假设值，所以拒绝零假设：如果建立了类似式(7-43)的置信区间，那么100个这样的置信区间中将有 95个包括真实的B_2值。我们不能说某一特定的区间(7-43)包括或不包括B_2的概率为95%。

勿庸赘言，可用两种不同的方法假设检验模型(7-32)中的其他任何一个系数。我们来看抵

1 附录7A.4给出的p值是双边检验的结果。要想得到单边检验的p值，只需将表中的数值二等分。

押贷款费用这个变量的系数，它在统计上是显著的吗？也即，它显著不为零吗？这依赖于我们如何建立零假设和备择假设。由于预期抵押贷款费用的系数为负，因此，需建立如下零假设和备择假设：

$$H_0: B_3=0，\ H_1: B_3<0 \tag{7-44}$$

即，用单边t检验。从t分布表可知，在5%显著水平下，单边t临界值为1.771(自由度为13)。式(7-32)回归结果中的t值为1.79，它超过了t临界值，所以能够拒绝零假设：抵押成本对抵押贷款无影响。[1]事实上，由式(7-32)回归结果中的p值，我们能在0.048(=0.092/2)的显著水平下拒绝零假设，这是获此t值1.79的精确的显著水平。[2]这个显著水平也是能够拒绝零假设的最低的显著水平。再一次的指出：如果能够得到p值，可用它来代替任意选择的显著水平α。

但是，如果备择假设为：$B_3 \neq 0$(也即双边检验)，那么就不能够拒绝零假设：抵押贷款费用对抵押贷款债务没有影响。因为，双边检验的t临界值为2.160，它比计算的t值1.79大。你会看到慎重地建立零假设与备择假设有多么重要。

7.8 对联合假设的检验

从式(7-32)回归结果可以看出：部分斜率系数b_2和b_3各自均在统计上是显著的；也即每个部分斜率系数均显著不为零。但是现考虑下面的零假设：

$$H_0：B_2=B_3=0 \tag{7-45}$$

这个零假设成为联合假设(joint hypothesis)，即B_2、B_3联合或同时为零(而不是各自的或单独的为零)。这个假设表明两个解释变量一起对应变量Y无影响；也就是说，抵押贷款债务独立于个人收水平以及抵押贷款费用。换句话说，X_2与X_3对Y无任何影响，等同于：

$$H_0：R^2=0 \tag{7-46}$$

也即，两个解释变量对应变量变化的解释比例为零(回忆R^2的定义)。因此，假设(7-45)与假设(7-46)是等价的。对这两个中任何一个假设进行检验称为对估计的总体回归线的显著性检验，即检验Y是否与X_2和X_3线性相关。

如何对(7-45)假设进行检验呢？这里有一个迷惑之处：既然B_2、B_3各自均显著不为零，那么它们一定也联合或集体显著不为零，也即能够拒绝(7-45)这个零假设。换句话说，既然个人收入和抵押贷款费用每一个都对抵押贷款数量有显著的影响，那么它们一起也一定会对抵押贷款有显著性影响。但是，这里需小心的是：在实践中的许多多元回归模型中，一个或多个解释变量各自对应变量没有影响，但集体却对应变量有影响，对于这一点我们将在第10章中结合多重共线性问题更详细地讨论。这意味着前面讨论的t检验虽然对于检验单个回归系数的统计显著性是有效的，但是对联合假设却是无效的。

那么如何对类似式(7-45)这样的假设进行检验呢？我们可用方差分析(analysis of variance，ANOVA)的技术来完成。下面介绍方差分析技术，首先看下面恒等式：

$$\text{TSS}=\text{ESS}+\text{RSS} \tag{7-33}$$

即，

$$\sum y_t^2 = b_2 \sum y_t x_{2t} + b_3 \sum y_t x_{3t} + \sum e_t^2 \tag{7-47}$$

---

1 与前面讲过的一样，如果检验统计量为负值，我们就考虑其绝对值。比如说，如果检验统计量的绝对值超过临界值，则将拒绝零假设。

2 (7-32)回归结果给出的p值是双边检验的结果。因此，对于单边检验，需将这个p值除以2。

式(7-47)将TSS分解为两个部分，一部分(ESS)由回归模型来解释，另一部分(RSS)不能由模型解释。对TSS的各个组成部分进行的研究称为方差分析。

正如第4章中所讲的，每一个平方和都与其自由度有关 —— 即计算的平方和是依据独立观察值的数目。于是上面的这些平方和的自由度为：

平 方 和	自 由 度
TSS	$n-1$(对所有模型，为什么？)
RSS	$n-3$(三变量模型)
ESS	2(三变量模型)*

* 求ESS自由度的一个简便方法是用TSS的自由度减RSS的自由度。

表7-3 三变量回归模型的方差分析表

方 差 来 源	平 方 和(SS)	自 由 度(d.f.)	MSS = SS/d.f.
来自回归(ESS)	$b_2\sum y_t x_{2t} + b_3\sum y_t x_{3t}$	2	$\left(b_2\sum y_t x_{2t} + b_3\sum y_t x_{3t}\right)/2$
来自残差(RSS)	$\sum e_t^2$	$n-3$	$\dfrac{\sum e_t^2}{n-3}$
总离差(TSS)	$\sum y_t^2$	$n-1$	

注：MSS＝平方和的平均值。

我们建立方差分析表，见表7-3。

现在，若满足CLRM基本假定(以及A7.6假定)，零假设为：H_0:B_2=B_3=0，可以证明变量：

$$F = \frac{\text{ESS/d.f.}}{\text{RSS/d.f.}} \tag{7-48}$$

$$= \frac{\text{被}X_2\text{和}X_3\text{解释的变动}}{\text{未被解释的变动}}$$

$$= \frac{\left(b_2\sum y_t x_{2t} + b_3\sum y_t x_{3t}\right)/2}{\sum e_t^2/(n-3)} \tag{7-49}$$

服从分子自由度为2，分母自由度为$n-3$的F分布。(参见第3章有关F分布的讨论及第4章有关F分布的使用)。一般地，如果回归模型有k个解释变量(包括截距)，则F值的分子自由度为$(k-1)$，分母自由度为$n-k$。[1]

如何利用式(7-49)给出的F值来检验联合假设：X_2和X_3对Y没有影响呢？我们可从式(7-49)中得到答案。如果分子比分母大，也即如果Y被回归解释的部分(即由X_2和X_3解释的Y的变动)比未被回归解释的部分大，则F值将大于1。因此，随着解释变量对应变量Y的变动的解释比例逐渐增大，F值也将逐渐增大。因此，F值越大，就越有理由拒绝零假设：两个(或多个)解释变量对应变量Y无影响。

当然，这种直观的原因可用假设检验的语言加以正规化。在第3章我们讲过，根据式(7-49)可计算出F值，并在所选显著水平下，将其与F临界值(分子自由度为2，分母自由度为$n-3$)作比较。如果计算的F值超过F临界值，则拒绝零假设：所有的解释变量同时为零。如果F值不超过F临界值，则不能拒绝零假设：解释变量对因变量无任何影响。

我们仍用抵押贷款的一例来说明。表7-4给出了对应表7-3的具体数值，该结果是用EVIEW

1 一个简单地记忆方法是：F值的分子自由度等于模型中偏斜率的个数，分母自由度等于n减去所需估计参数的个数(即，偏斜率系数加上截距)。

软件计算得到(见附录7A.4)。[1]从表7-4的计算结果可知，估计的F值为608.918 5，约为609，在零假设$B_2=B_3=0$，若满足古典线性回归模型的基本假定，那么，F值服从分子自由度为2，分母自由度为3的F分布。

如果零假设为真，则得此F值大于或等于609的概率是多少呢？计算机计算结果表明，获此F值的p值为0.000 00，几乎为零。因此，能够拒绝零假设：收入和抵押贷款费用一起对抵押贷款债务没有影响。更肯定地说，个人收入和抵押贷款费用确实影响了抵押贷款债务。[2]

表7-4　抵押贷款债务一例的方差分析表

方 差 来 源	平方和(SS)	自由度(d.f.)	MSS=SS/d.f.
来自回归(ESS)	19020058	2	19 020 258/2=9 510 029
来自残差(RSS)	20 303 3	13	203 033/2=15 618
总离差(TSS)	19 233 091	15	
F=9 510 029/15618=608.916①			

① 数值已经经过四舍五入，因此这个F值与计算机输出结果略有出入。

在我们这个例子中，不仅拒绝假设：B_2和B_3分别是统计非显著的，而且拒绝假设：B_2和B_3联合地是统计非显著的。然而，这种情况并不总是发生。我们将会遇到不是所有的解释变量各自都对应变量有影响(也就是，有的t值可能是统计不显著的)但是它们却联合地对应变量有影响(即F检验将拒绝零假设：所有斜率同时为零)的情况。我们将会看到，当存在多重共线性时，上述情况就会发生。有关多重共线性问题更详细的讨论见第十章。

F与R^2之间的重要关系

判定系数R^2与方差分析中用到的F值之间有如下重要关系：(参见习题7.16)

$$F=\frac{R^2/(k-1)}{(1-R^2)/(n-k)} \tag{7-50}$$

式中　n为观察值的个数；　k为包括截距在内的解释变量的个数。

式(7-50)表明了F与R^2之间的关系。这两个统计量同方向变动。当$R^2=0$(即Y与解释变量X不相关)时，F为0。R^2值越大，F值也越大。当R^2取其极限值1时，F值为无穷大。

因此，前面讨论过的F检验(用于度量总体回归直线的显著性)也可用于检验R^2的统计显著性——即R^2是否显著不为零。换句话说，检验零假设(7-45)与检验零假设(总体的)R^2为零是等价的。

用R^2的形式进行F检验的一个优点是便于计算。我们仅仅需知道R^2值即可，而许多统计软件都给出R^2的值。因此，对总体回归方程显著性的F检验可以用R^2的形式，见式(7-50)。相应的，表7-3的方差分析表也可等价地表示为表7-5：

表7-5　用R^2形式表示的方差分析表

方 差 来 源	平方和(SS)	自由度(d.f.)	MSS=SS/d.f.
来自回归(ESS)	$R^2\sum y_t^2$	2	$\left(R^2\sum y_t^2\right)/2$
来自残差(RSS)	$(1-R^2)\sum y_t^2$	$n-3$	$\dfrac{(1-R^2)\left(\sum y_t^2\right)}{(n-3)}$
总离差(TSS)	$\sum y_t^2$	$n-1$	

注：1.在计算F值时，无需将R^2和$(1-R^2)$分别乘以$\sum y_t^2$，因为$\sum y_t^2$可被约去，见式(7-50)。
2.在k个变量的模型中，自由度分别为$(k-1)$和$(n-k)$。

1　与其他统计软件不同，EVIEWS虽然给出了F值，但不提供方差分析表。但是从附录7A.4给出的结果，不难求得方差分析表，因为，我们可用给出的Y的标准方差来计算TSS。我们还有RSS数据，因此，可计算出ESS。当然，各个不同的自由度也是已知的。

2　如果选择$\alpha=1\%$，则F临界值为6.70（分子自由度为2，分母自由度为13），显然，计算的F值609几乎是F临界值6.70的91倍。

在这个例子中，R^2＝0.989。因此，根据式(7-50)：

$$F = \frac{0.9894/2}{(1-0.9\,894)/1\,3} = 606.78 \tag{7-51}$$

这里的F值与附录7A.4中计算机输出结果几乎相同，只是舍入误差略有不同。

按照表7-5的形式建立方差分析表，留给读者实践。

7.9 从多元回归模型到双变量模型：设定误差

我们再来看表7-1提供的抵押贷款债务的数据。假定某人认为住房对他很重要以至于认为抵押贷款费用不像个人收入那么重要。因此，他或她决定把抵押贷款费用这一解释变量(X_3)从回归方程(7-2)中略去，仅仅简单地做Y(抵押贷款债务)对X(个人收入)的回归，得到如下结果：

$$\begin{aligned} \hat{Y}_t &= -861.7+0.929\,3X_{2t} \\ \text{se} &= (122.5)\quad (0.028\,7) \qquad R^2=0.987 \\ t &= (-7.03)\quad (32.38) \qquad \text{校正的}R^2=0.986 \\ p\text{值} &= (0.000)^*\quad (0.000)^* \end{aligned} \tag{7-52}$$

*表示值很小。

如果把双变量回归结果(7-52)与三变量回归结果(7-32)作比较，你会发现有几个不同：

(1) 虽然抵押贷款债务仍与个人收入高度相关，但是两个回归方程中收入变量的系数不同，分别为0.825 8[式(7-32)]，0.929 3[式(7-52)]，而且这种差别在统计上是显著的。(你能证明吗？)

(2) 两回归模型中的截距相差很大。

(3) 两回归模型中的R^2一个为0.987，另一个为0.989，看似差别不大。这是完全可能的，而且随后我们还将看到,这种有限的差别可能在统计上还是显著的，也即显著不为零。

两个回归结果为什么会不同呢?别忘了在式(7-32)中，我们推导个人收入对抵押贷款债务的影响时,是在假设抵押贷款费用是常数的条件下;而在式(7-52)中，只是简单地略去了抵押贷款费用这个变量。换句话说,式(7-32)中的个收入对抵押贷款债务的影响是净影响或净效果,而在式(7-52)中的抵押贷款费用的影响并未略掉。因而 式(7-5)中的个人收入系数反应了总效果——直接的个人收入效果与间接的抵押贷款费用的效果。两个回归结果的这种差异性很好地反映了偏回归系数的“偏”的含义。

在对式(7-32)回归结果的讨论中，我们发现个人收入和抵押贷款费用无论是各自地，还是是集体地都对抵押贷款债务有重要影响。因此，从式(7-32)的回归模型中将抵押贷款变量略去，会导致(模型的)设定偏差(model)specification)或设定误差(specification error)，更具体说是从模型中略去重要变量的设定误差。

在第14章中将详细讨论模型的设定误差，但这里需要特别指出的是：在依据经验建立模型时需格外仔细。在建立模型时，要以经济理论为依据并充分利用以往的工作经验。一旦建立起模型，就不要任意地从模型中删除某个解释变量。

7.10 两个不同的R^2的比较：校正的判定系数

检查双变量回归模型[式(7-52)]与三变量回归模型[式(7-32)]的R^2值，我们注意到前一个方程的R^2值(0.987)比后一个方程的R^2值(0.989)小。结果总是这样的吗？是的！判定系数R^2的一个重要性质就是模型中的解释变量的个数越多，R^2值就越大。那么，若想用更大的比例解释应变量的变动，那么仅仅需要不断地增加解释变量的个数就可以了！

但是，不能完全照搬这个“建议”。这是因为在R^2的定义(ESS/TSS)中并没有考虑到自由度。注意：在一个有k个变量的模型中(包括截距)，ESS的自由度为$(k-1)$。因此，如果模型中有5个解释变量(包括截距)，则ESS的自由度为4，如果模型有10个解释变量(包括截距)，则ESS的自由度为9，但是R^2的计算公式并未考虑不同模型自由度的不同[TSS的自由度总为$(n-1)$(为什么？)]。因此，比较相同应变量但不同个数解释变量的两回归模型的样本判定系数R^2，就像是拿苹果与桔子比。

因而，我们需要这样一个拟合优度的度量指标，它能根据模型中解释变量的个数进行调整。定义校正的判定系数$\bar{R}^2$(adjusted R^2)如下：(参见附录7A.3)

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \tag{7-53}$$

这里需要指出的是前面讨论的判定系数R^2也称为非校正的判定系数。

校正的判定系数有如下性质：

(1) 若$k>1$，则$\bar{R}^2 \leqslant R^2$。即，随着模型中解释变量的增加，校正判定系数越来越小于非校正判定系数R^2，这似乎是对增加解释变量的“惩罚”。

(2) 虽然非校正的判定系数$\bar{R}^2$总为正，但校正的判定系数$\bar{R}^2$可能为负。例如，若回归模型中，$k=3$，$n=30$，$R^2=0.06$，则$\bar{R}^2$为$-0.009\ 6$。

目前，许多统计软件都可以计算R^2和$\bar{R}^2$。校正判定系数可以使我们对同应变量不同解释变量(个数不同)的两回归模型作比较。[1]

在抵押贷款债务一例中，校正判定系数$\bar{R}^2$为0.987 8，比非校正的判定系数$R^2$0.989 4略小。

7.11 什么时候增加新的解释变量

在实际中，为了解释某一现象，研究者往往面对如何取舍若干解释变量的问题。通常的做法是：只要校正判定系数$\bar{R}^2$值增加(即使$\bar{R}^2$值可能小于非校正判定系数R^2的值)，就可以增加解释变量。但是什么时候校正的判定系数$\bar{R}^2$值开始增加呢？可以证明：如果增加变量的系数的$|t|$值大于1，$\bar{R}^2$就会增加，这里的t值是在零假设：总体的系数为零下计算得到的。

在抵押贷款债务一例中，已经知道$\bar{R}^2=0.986\ 0$(双变量模型)，0.987 8(三变量模型)。校正判定系数值增加了，但这并没有什么可奇怪的，因为变量X_3的t值[见式(7-32)]超过1。因此，可将X_3增加到模型中去。同时，X_3的t值不仅比1大，而且根据单边检验的结果，它还在统计显著的。[2]

7.12 回归模型的结构稳定性检验：Chow检验

当回归模型涉及时间序列数据时，应变量Y与解释变量X之间可能会发生结构变化。有些时候，结构的变化可能是由于内部因素(例如，在1973年、1979年以及1990～1991年的海湾战争期间，由OPEC石油卡特尔组织发起的石油禁运)，有些时候，结构的变化可能是由于政策的变化(例如，1973年间，由固定汇率制转换为浮动汇率制)所导致的，或是由于国会采取的某些行动(里根政府推行的新税收政策)，或是由于其他的原因。

如何发现模型中确实发生了结构变化呢？具体地,我们考虑表7-6提供的数据,它给出了

1 在第8章中将会看到：如果两个回归模型应变量不同，则不能直接比较它们的判定系数R^2和校正的判定系数$\bar{R}^2$。

2 但是无论某t值是否在统计上是显著的，只要增加变量的系数的$|t|$值大于1，校正的判定系数$\bar{R}^2$就会增加。

1970~1995年间美国个人可支配收入与个人储蓄的数据，假定我们想要估计一个简单的储蓄方程,即个人储蓄(Y)对个人可支配收入(X)的变化。根据数据，可以得到 Y关于X的OLS回归方程。但是如果这样做，就隐含地假定了在 26年间Y与X的关系没有发生太大的变化。这可能是一个过于理想的假定，比如，在1982年，美国遭受到了和平期间最严重的经济衰退，当年的城市失业率高达9.7%，是自1948年以来失业率最高的一年。类似这种事件会扰乱收入和储蓄之间的关系。为了观察这种情况是否发生，我们把数据分为两个时期： 1970~1981年和1982~995年。得到三种可能的回归方程：

时期：1970～1995年 $Y_t=A+BX_t+u_{1t}$ $n=26$ (7-54)

时期：1970～1981年 $Y_t=A_1+B_2X_t+u_{2t}$ $n_1=12$ (7-55)

时期：1982～1995年 $Y_t=B_1+B_2X_t+u_{3t}$ $n_3=14$ (7-56)

(注：$n=n_1+n_2$)

式(7-54)回归假定了两个时期没有差别，因此，估计的是整个时期的 Y与X之间的关系。(7-55)回归与(7-56)回归假定了Y与X的关系在两个时期是不同的。

利用表中数据，得到如下回归结果：

$$\hat{Y}_t = 62.422\,6 + 0.037\,6X_t \quad (7\text{-}54a)$$

$t=$ (4.891 7) (8.893 7) $R^2=0.767\,2$ d.f.=24 $S_1=232\,48.30$

$$\hat{Y}_t = 1.106\,1 + 0.080\,3X_t \quad (7\text{-}55a)$$

$t=$ (0.087 3) (9.601 5) $R^2=0.902\,1$ d.f.=10 $S_2=178\,5.032$

$$\hat{Y}_t = 153.494\,7 + 0.014\,8X_t \quad (7\text{-}56a)$$

$t=$ (4.692 2) (1.770 7) $R^2=0.207\,1$ d.f.=12 $S_3=100\,05.22$

注：S_1，S_2，S_3分别表示各个回归方程的残差平方和(RSS)，这里仅给出了t值。

表7-6 美国个人可支配收入与个人储蓄

年 份	个人储蓄	个人可支配收入	年 份	个人储蓄	个人可支配收入
1970	61.000 00	727.100 0	1983	167.000 0	2 522.400
1971	68.600 00	790.200 0	1984	235.700 0	2 810.000
1972	63.600 00	855.300 0	1985	206.200 0	3 002.000
1973	89.600 00	965.000 0	1986	196.500 0	3 187.600
1974	97.600 00	1 054.200	1987	168.400 0	3 363.100
1975	104.400 00	1 159.200	1988	189.100 0	3 640.800
1976	96.400 00	1 276.000	1989	187.800 0	3 894.500
1977	92.500 00	1 401.400	1990	208.700 0	4 166.800
1978	112.600 0	1 580.100	1991	246.400 0	4 343.700
1979	130.100 0	1 769.500	1992	272.600 0	4 613.700
1980	161.800 0	1 973.300	1993	214.400 0	4 790.200
1981	199.100 0	2 200.200	1994	189.400 0	5 021.700
1982	205.200 0	2 374.300	1995	249.300 0	5 320.800

资料来源：《总统经济报告》(*Economic Report of the President*)，1997，所有数据的单位均为亿美元，数据摘自表B-28，第332页。

观察上述回归结果，可以看出收入和储蓄间的关系在不同时期有所不同。这表明合并的回归(7-54a)或许不是很准确(即拟合所有 26个观察值，得到一个共同的回归方程，而不考虑两个时期可能有差别)。两个不同时期回归结果的差异可能是由于结构的变化，这种结构的变化或许是由于截距(B_1和A_1在统计可能是不同的)或斜率(B_2，A_2在统计可能是不同的)的不同，或是斜率与截距都不同。我们如何判定呢？

这就要用到Chow检验。[1]Chow检验假定：

(1) u_{2t}~N(0,σ^2)，u_{3t}~N(0,σ^2)即(7-55)和(7-66)回归方程中的随机误差项服从同方差正态分布。

(2) 两随机误差项u_{1t}和u_{2t}相互独立。

Chow检验步骤如下：

(1) 估计(7-54)回归方程，如果不存在结构变化，则是一个准确的回归方程。求当自由度为(n_1+n_2-k)时的RSS及S_1，其中，k是估计参数的个数。在这个例子中，S_1=23 248.30。我们称S_1为限制的残差平方和(RSS_R)，因为它是在假定$A_1=B_1$，$A_2=B_2$的条件下(也即回归方程(7-55)与(7-56)无差别)得到的。

(2) 估计(7-55)回归方程，并求在自由度为(n_1-k)时的RSS，S_2。在本例中，S_2=1 785.032。

(3) 估计(7-56)回归方程，并求在自由度为(n_2-k)时的RSS，S_3，在本例中，S_3=10 005.22。

(4) 由于两组样本相互独立，所以可将S_2与S_3相加，得到非限制残差平方和(RSS_{UR})

$$RSS_{UR}=S_2+S_3，\ d.f.=n_1+n_2-2k$$

在本例中，$RSS_{UR}=S_2+S_3$=11 790.252

(5) Chow检验的思想是：如果模型中确实不存在结构变化［也即，(7-55)回归方程与(7-56)相同]，那么RSS_R与RSS_{UR}不是统计不同的。因此，如果得到如下的F值

$$F=\frac{(RSS_R-RSS_{UR})/k}{RSS_{UR}/(n_1+n_2-2k)} \tag{7-57}$$

那么根据Chow检验，在零假设：两个回归方程(7-55)与(7-66)统计相同的(即不存在结构的变化)下，式(7-57)给出的F值服从分子自由度为k，分母自由度为(n_1+n_2-2k)的F分布。

(6) 因此，如果计算的F值没有超过F临界值，那么不能拒绝参数是稳定性(即不存在结构变动)的零假设。在这种情况下，我们可以使用(7-54)回归方程。相反地，如果计算的F值超过了F临界值，则拒绝参数是稳定的假设，并得出结论：回归方程(7-55)与(7-56)是不同的，在这种情况下，回归方程(7-54)的参数值是不确定的。

再来看这个例子，我们发现：

$$F=\frac{(2\,324.80-11790.252)/22}{(11790.252)/22} \tag{7-58}$$
$$=10.69$$

根据F分布表，可得在1%的显著水平下，F临界值为7.72(分子自由度为2，分母自由度为22)。因此，得到F值大于等于10.69的概率小于1%：精确地说，p值仅为0.00 057。

因此，得出结论：储蓄-收入回归方程(7-55)与(7-60)是不同的，即储蓄函数经历了一个结构的变动。这可能由于里根时期的经济政策或是其他别的因素。那么显然，(7-54)回归就没有什么意义了。

在运用Chow检验时，需要注意以下的一些限制条件：

(1) 必须满足上面讲到的假定条件。例如，必须确认回归方程(7-55)与(7-56)的误差项是否相同。在第11章中，将讨论如何处理异方差情况。

(2) Chow检验的结果仅仅告诉我们两回归方程(7-55)和(7-56)是否不同，而无法得知导致这种差异的原因。在第9章关于虚拟变量的讨论中，将对这个问题作出回答。

---

1 参见Gregory C.Chow的《两线性方程回归系数相同性检验》(Tests of Equality Between Sets of Coefficients in Two Linear Regressions)，见《经济计量学家》(Econometrica)，第28卷，第3期，1960年，第591-605页。

(3) Chow检验假定知道结构发生变化的时间点。在本例中，是在 1982年。但是，如果不清楚这个信息，我们需要用其它的方法。[1]

在本书后面章节将会看到，(7-57)回归中给出的F检验有着更广泛的用途。

7.13 实例

为了概括本章，我们考虑几个多元回归的例子。旨在向读者介绍多元回归在实际中的应用。

例7.1 税收政策会影响公司资本结构吗

Pozdena估计了下面的回归方程：[2]

其中，

Y＝杠杆利率（＝债务/产权） X_2＝公司税率 X_3＝个人税率 X_4＝资本利得税率 X_5＝非债务的避税 X_6＝通货膨胀率

表7-7 制造业中杠杆作用

解释变量	系数(括号内的是t值)	解释变量	系数(括号内的是t值)
公司税率	2.4	个人税率	−1.2
	(10.5)		(4.8)
资本利得税率	0.3	非债务避税	−2.4
	(1.3)		(−4.8)
通货膨胀率	1.4		
	(3.0)		
n=48(观察值个数)	R^2=0.87	$\bar{R}^2$ = 0.85	

资料来源：Randall Johnston Pozdena, 《税收政策与企业资本结构》 (Tax Policy and Corporate Capital Structure)，《经济评论》(Economic Review)，旧金山联邦储备银行，1987年秋，表1，第45页。

注：1.作者并未给出估计的截距值。

2.校正判定系数是根据式(7-53)计算得到的。

3.各系数的标准差可由系数除以其t值得到(例如，2.4/10.5=0.228 6即是公司税率系数的标准差。)

根据经济理论可知,系数B_2、B_4、B_6为正，B_3、B_5为负。[3]根据美国1935～1982年间制造业的数据,Pozdena以列表式(表7-7)给出了OLS结果。(有的时候回归结果用列表形式给出,以便阅读。)

对回归结果的讨论

第一个需要指出的事实是：上述回归结果中，所有系数的符号与经济理论预期相符。举个例子，公司税率对杠杆利率有正的影响 —— 假定其它条件保持不变，公司税率每增加一个百分点，平均而言，杠杆利率(即债务产权的比率)将上升 2.4个百分点。同样地，在其它条件不变时，如果通货膨胀率每上升每一个百分点，平均而言，杠杆利率将上升 1.4个百分点。其它的偏回归系数可类似地解释。

1 详细的讨论参见Willam H.Greene的《经济计量分析》(*Econometric Analysis*)，第3版，Prentice-Hall,新泽西，1997年，第355页。

2 资料来源：Randall Johnston Pozdena的《税收政策与企业资本结构》(*Tax Policy and Corporate Capital Structure*)，《经济评论》(Economic Review)，旧金山联邦储备银行，1987年秋，第37~51页。

3 参见Pozdena的文章（脚注14）对各个系数预期符号的理论分析。需要指出的是，美国债务利率是税后利率，但是收入利率是税前利率，这就是为什么公司宁愿负债也不愿将资本投资到新企业的原因之一。

在每一个偏回归系数的下面给出了在零假设(每一个总体偏回归系数分别为零)下的 t值。所以我们能够很容易地检验这个零假设和(双边的)备择假设：每一个真实的总体系数不为零。因此，在这个例子中，我们用双边 t检验。自由度为42。(注：表7-7并未给出截距值，虽然它也被估计了)。如果选择显著水平 $\alpha=5\%$，则双边t检验值约为2.021(d.f.=40)。(注：虽然t分布表并未给出自由度为42时的精确的t值，但这已是一个很好的近似值)，如果选择显著水平 $\alpha=1\%$，则双边的t检验值为2.704，看表7-7给出的t值，在1%的显著水平下，显著不为零。在1%或5%的显著水平下，资本所得税的系数都是不显著的。因此，除此变量外，能够拒绝各个零假设：每一个偏回归系数为零。换句话说，除了一个解释变量之外，所有其他的解释变量各自均对债务/产权比例有影响。这里指出：如果一个估计的系数在 1%显著水平下是统计显著的，则它在 5%的水平下同样是统计显著的，但是，反之，则不成立。

整个估计的回归直线的显著性如何呢？即，能否接受零假设：所有的偏斜率同时为零或，等价地，$R^2=0$？这个假设很容易用式(7-50)来检验，在此例中，

$$F=\frac{R^2/(k-1)}{(1-R^2)/(n-k)}=\frac{0.87/5}{0.13/42}=56.22 \tag{7-60}$$

它服从自由度为5和42的F分布，如果$\alpha=0.05$，从F分布表可知，F临界值为2.45。(d.f.=5,40，因为F分布表没有分母自由度为40时相应的F值)。在$\alpha=0.01$，相应的F临界值为3.51，计算的F值56远远大于上面任何一个F临界值。因此，拒绝零假设：所有的偏斜率同时为零，或等价地，$R^2=0$，集体地，所有5个变量都影响应变量。个别地，只有 4个变量对应变量有影响，例 7.1再一次强调了：(单个)t检验与(联合)F检验完全不同。[1]

例7.2 牙买加对进口的需求

为了解释牙买加对进口的需求，J.Gafar[2]根据19年的数据得到下面的回归结果：

$$\hat{Y}_t=-58.9+0.20X_{2t}-0.10X_{3t} \tag{7-61}$$

se= (0.0 092) (0.084) $R^2=0.96$

t= (21.74) (−1.1904) $\bar{R}^2=0.955$

其中，

Y＝进口量

X_2＝个人消费支出

X_3＝进口价格/国内价格

经济理论表明Y与X_2之间正相关，Y与X_3之间负相关，这与回归结果相符。在5%的显著水平上，X_2的回归系数在统计上是显著的，但 X_3的系数却不是。但是由于（X_3的）t的绝对值大于1，因而如果将X_3从模型中略去，校正的判定系数 $\bar{R}^2$将减少（为什么？）。联合地，X_2和X_3可以解释牙买加进口数量96%的变化。

---

1 我们知道在双变量线性回归模型中，在两个变量之间存在着如下关联：

$$t_k^2=F_{1,k}$$

也就是说，具有k个自由度的t值的平方等于分子具有1个自由度并且分母自由度为k的F值（见式(3-18)）。

2 参见J.Gafar的《发展中经济的贬值与国际收支调整：以牙买加为例》(Devaluation and the Balance of Payments Adjustment in a Developing Economy:An Analysis Relating to Jamaica)，《应用经济学》(*Applied Economics*)，第29卷，1981，第151-165页。符号略有变动。

例7.3 英国对酒精饮料的需求

为了解释英国对酒精饮料的需求，T.McGuinness[1]根据20年的年数据得到了下面的回归结果：

$$\begin{aligned}\widehat{Y}_t &= -0.014 - 0.354X_{2t} + 0.001\,8X_{3t} + 0.657X_{4t} + 0.005\,9X_{5t}\\ \text{se} &= (0.012)\quad (0.2688)\quad (0.0005)\quad (0.266)\quad (0.003\,4)\\ t &= (-1.16)\quad (1.32)\quad (3.39)\quad (2.47)\quad (1.73)\\ R^2 &= 0.689\end{aligned} \tag{7-62}$$

其中，

Y＝每一个成年人纯酒精消费的年变化

X_2=酒精饮料的真实价格指数的年变化

X_3=个人真实的可支配收入的年变化

$X_4 = \dfrac{\text{许可证颁发数量年变化量}}{\text{成年人人口}}$

X_5=在酒精饮料上的广告支出费用的年变化

理论表明，除了变量X_2以外，所有解释变量都与应变量Y正相关。这也与回归结果相一致，虽然有的回归系数是统计不显著的。在5%的显著水平下，（自由度为15，为什么？）t临界值为1.73（单边）和2.131（双边）。考虑变量广告支出X_5的系数，由于预期广告支出与对酒精饮料的需求正相关（不然的话，这对广告行业就是一个坏消息），于是可以建立假设：H_0:B_5=0,H_1:B_5>0。因此我们可以用单边t检验。在5%的显著水平下，计算得到的t值近似统计显著的。

留给读者计算F值，并检验假设：所有的偏斜率系数同时为零。

例7.4 城市劳动参与率，失业率以及平均小时工资

在第一章中，我们给出了(1-5)回归方程，但并未对回归结果的统计显著性进行讨论。现在我们已经具备了必要的工具来讨论这个回归结果。完整的回归结果如下：

$$\begin{aligned}\widehat{CLFPR}_t &= 97.936 - 0.446\,3CUNR_t - 3.858\,9AHE82_t\\ \text{se} &= (5.513)\quad (0.0949)\quad (0.7552)\\ t &= (17.16)\quad (-4.70)\quad (-5.11)\\ p\text{值} &= (0.000)^*\quad (0.000)^*\quad (0.000)^*\qquad R^2 = 0.8\\ & \qquad F = 40.88\qquad\qquad \text{校正的}R^2 = 0.8\end{aligned} \tag{7-63}$$

*表示值很小。

回归结果表明，每个估计的回归系数各自均统计显著的，因为计算出的p值很小。也就是说，每个系数显著不为零。集体的，CUNR和AHE也是统计显著的，因为获得F值为40(自由度为2和14)的概率非常小。

与预期相同，城市失业率对城市劳动参与率有负的影响，表明失业工人效果占统治地位。隐含着的经济理论已经在第1章中解释过。AHE82的系数为负，表明在收入效应和替代效应中，前者占统治地位。

1 参见T.McGninness的《英国对酒精饮料需求的经济计量分析》(An Econometric Analysis of Total Demand of Alcoholic Beverages in the United Kingdom)，《工业经济》(Journal of Industrial Economics)，1980，第29卷，第85-109页，符号略有变动。

7.14 小结

本章我们讨论了最简单的多元回归模型，三变量线性回归模型 —— 一个应变量，两个解释变量。虽然在许多方面，三变量模型是双变量线性回归模型的直接扩展，但是通过三变量模型，我们介绍了一些新的概念，比如偏回归系数，校正的和非校正的多元判定系数，多重共线性等。

就多元回归参数估计而言，我们仍然是在普通最小二乘估计的框架下进行参数估计的。与双变量模型相同，多元回归的 OLS估计量也具有很好的统计性质 —— 最优线性无偏估计(BLUE)的高斯-马尔科夫性质。

在扰动项服从均值为零，方差为 σ^2的假定下，与双变量模型相同，每个估计的系数都服从正态分布(均值等于真实的总体值，方差见文中给出的计算公式)。不幸的是，在具体实践中，σ^2是未知的，需要估计求出，而这个未知参数的 OLS估计量是 $\hat{\sigma}^2$ 。如果用 $\hat{\sigma}^2$ 代替 σ^2，那么，与双变量模型相同，每个估计的回归系数服从 t分布，而不是正态分布。

每个多元回归系数服从自由度为 $(n-k)$的t分布(其中k是包括截距在内的需要估计的参数的个数)，意味着我们能够用 t分布对每个多元回归系数的统计显著性进行检验假设。我们可以用两种不同的检验方法 —— 显著性检验法和置信区间法进行假设检验。在这方面，多元回归模型与双变量模型并没有什么不同，只是自由度有所限制。

但是，当检验假设：所有的偏斜率系数同时为零时，前面提到的单个 t检验就没有用了。这里需要用方差分析表及 F检验。同时，检验假设：所有的偏斜率系数同时为零与检验假设：样本判定系数 R^2为零是等价的。因此， F检验也可以由于后者。我们还证明了 F检验(Chow 检验)是如何用于检查回归模型的结构稳定性的。

本章我们通过一些具体的数字例子介绍了这些概念并讨论了它们在经济实践中的具体运用。

习题

7.1 解释概念

(a) 偏回归系数 (b) 多元判定系数，R^2 (c) 完全共线性 (d) 完全的多重共线性 (e) 单个假设检验 (f) 联合假设检验 (g) 校正判定系数 $\bar{R}^2$

7.2 按步聚解释下类过程。

(a) 对单个多元回归系数的显著性检验 (b) 对所有的部分斜率系数的显著性检验

7.3 判断正误并说明理由。

(a) 校正的判定系数和非校正的判定系数仅当非校正判定系数为 1时才相等。 (b) 判定所有解释变量是否对应变量有显著影响的方法是看是否每个解释变量都是显著的 t统计量；如果不是，则解释变量整体是统计不显著的。 (c) 当R^2=1, F=0;当R^2=0,F=∞。 (d) 当自由度大于120时，在5%显著水平下，(双边检验的) t临界值与在5%显著水平下的(标准正态变量) Z临界值相同，均为1.96。 (e) 在模型$Y_i=B_1+B_2X_{2i}+B_3X_{3i}+u_i$中，如果$X_2$和$X_3$负相关且$B_3>0$,则从模型中略去解释变量$X_3$将使$b_{12}$的值减小[也即,$E(b_{12})<(B_2)$]。其中$b_{12}$是$Y$仅对$X_2$的回归方程中的斜率系数。(f) 当我们说估计的回归系数在统计上是显著的，意思是说它显著不为 1。 (g) 要计算t临界值，仅仅需知道自由度。 (h) 整个多元回归模型在统计上是显著的意味着模型中任何一个单独的变量均是统计显著的。 (j) 就估计和假设检验而言，单方程回归与多元回归没有什么区别。(k) 无论模型中包括多少个解释变量，总离差平方和的自由度总为($n-1$)。

7.4 求下列情况下 $\hat{\sigma}^2$ 的。

(a) $\sum e_i^2$ =800，n=25，k=4(包括截距) (b) $\sum e_i^2$ =1200，n=14，k=3(不包括截距)

7.5 求下列情况的t临界值

自由度(d.f.)	显著水平(2%)	H_1
12	5	双边
20	1	右边
30	5	左边
200	5	双边

7.6 求下列情况的F临界值。

分子自由度	分母自由度	显著水平(%)
5	5	5
4	19	1
20	200	5

7.7 已知下列数据：

Y	X_2	X_3
1	1	2
3	2	1
8	3	-3

根据上表数据，估计下列回归方程：

(a) $Y_i=A_1+A_2X_{2i}+u_i$ (b) $Y_i=C_1+C_3X_{3i}+u_i$ (c) $Y_i=B_1+B_2X_{2i}+B_3X_{3i}+u_i$

注：不必担心估计标准差。

(1) $A_2=B_2$？为什么？

(2) $C_3=B_3$？为什么？

从这个习题中，你能得出什么样的结论？

7.8 下面给出依据15个观察值得计算得到的数据：

$$\bar{Y}=367.693；\ \bar{X}_2=402.760；\ \bar{X}_3=8.0；\ \sum y_i^2=66\ 042.269$$

$$\sum x_{2i}^2=84\ 855.096；\ \sum x_{3i}^2=280.0；\ \sum y_i x_{2i}=74\ 778.346$$

$$\sum y_i x_{3i}=4\ 250.9；\ \sum x_{2i}x_{3i}=4\ 796.0$$

小写字母代表了各值与其样本均值的离差。

(a) 估计三个多元回归系数。 (b) 估计它们的标准差。 (c) 求R^2与$\bar{R}^2$？ (d) 估计B_2，$B_3$95%的置信区间。 (e) 在$\alpha=5\%$下，检验估计的每个回归系数的统计显著性(双边检验)。 (f) 检验在$\alpha=5\%$下所有的部分系数都为零。给出方差分析表。

7.9 下表给出了三变量模型的回归的结果：

方差来源	平方和(SS)	自由度(d.f.)	平方和的均值(MSS)
来自回归(ESS)	65 965	—	—
来自残差(RSS)	—	—	—
总离差(TSS)	66.042	14	

(a) 样本容量是多少？

(b) 求RSS？

(c) ESS与RSS的自由度各是多少？

(d) 求R^2与$\bar{R}^2$？

(e) 检验假设：X_2和X_3对Y的无影响。你用什么假设检验？为什么？

(f) 根据以上信息，你能否确定X_2和X_3各自对Y的贡献吗？

7.10 以R^2形式重写习题7.9中的方差分析表。

7.11 为了确定对空调价格的影响因素，B.T. katchford[1]根据19个样本数据得到回归结果如下：

$$\hat{Y}_2 = -68.26 + 0.023X_{2i} + 19.729X_{3i} + 7.653X_{4i} \qquad R^2 = 0.84$$

se= (0.005) (8.992) (3.082)

式中 Y——空调的价格/美元

X_2——空调的BTO比率

X_3——能量效率

X_4——设定数

(a) 解释回归结果。 (b) 该回归结果有经济意义吗？ (c) 在显著水平$\alpha=5\%$下，检验零假设：BTU比率对空调的价格无影响，备择假设检设：BTU比率对价格有正向影响。 (d)你会接受零假设：三个解释变量在很大程度上解释了空调价格的变动吗？详细写出计算过程。

7.12 根据美国1965－Ⅰ Q至1983-Ⅱ Q数据(n=26)，James Doti与Esmael Adibi[2]得到下面的回归方程用以解释美国的个人的消费支出(PCE)：

$$\hat{Y}_t = -10.96 + 0.93X_{2t} - 2.09X_{3t}$$

t=(－3.33) (249.06) (－3.09) $R^2=0.999\,6$

$F=93\,753.7$

式中 Y——个人消费支出/亿美元

X_2——可支配(税后)收入/亿美元

X_3——银行支付的主要利率(%)

(a) 求边际消费倾向(MPC)？——每额外增加1美元个人可支配收入所增加的消费支出的数量。 (b) MPC显著不为1吗？给出检验过程。 (c) 模型中包括主要利率变量的理论基础是什么？先验地,你预期这个变量的符号为负吗？ (d) b_3显著不为零吗？ (e) 检验假设$R^2=0$。(f) 计算每个系数的标准差。

7.13 对本章例7.2，检验假设：X_2和X_3一起对Y无影响。你用什么检验？在此检验下有哪些假定条件？

7.14 表7-8给出1980～1996年美国的城市劳动参与率、失业率等数据。

(a) 建立一个合适的回归模型解释城市男性劳动力参与率与城市男性失业率及真实的平均小时工资之间的关系。 (b)重服(a)过程，但此时的变量为女性城市劳动参与率。 (c)重复(a)过程，但此时用当前平均小时工资。 (d)重复(b)过程，但此时用当前平均小时工资。(e)如果 (a)和 (c)的回归结果不同，你如何解释？ (f)如果 (b)和 (d)的回归结果不同，你如何使回归结果合理化？

表7-8 劳动力参与数据

年份	CLFPRM①	CLFPRF②	UNRM③	UNRF④	AHE82⑤	AHE⑥
1980	77.4	51.5	6.9	7.4	7.78	6.66
1981	77.0	52.1	7.4	7.9	7.69	7.25
1982	76.6	52.6	9.9	9.4	7.68	7.68
1983	76.4	53.9	9.9	9.2	7.79	8.02
1984	76.4	53.6	7.4	7.6	7.80	8.32
1985	76.3	54.5	7.0	7.4	7.77	8.57
1986	76.3	55.3	6.9	7.1	7.81	8.76
1987	76.2	56.0	6.2	6.2	7.73	8.98
1988	76.2	56.6	5.5	5.6	7.69	9.28

1 参见B.T. katchford，(The Value of Information for Selected Appliance)，(Journal of Marketing Research)，第17卷，1980年，第14-25页。符号略有变动。

2 参见James Doti与Esmael Adibi的《经济计量分析：运用方法》(*Econometric Analysis:An Applications Approach*)，Pretice-Hall,Englewood,N.J.,1988年，第188页。符号略有改动。

(续)

年份	CLFPRM①	CLFPRF②	UNRM③	UNRF④	AHE82⑤	AHE⑥
1989	76.4	57.4	5.2	5.4	7.64	9.66
1990	76.4	57.5	5.7	5.5	7.52	10.01
1991	75.8	57.4	7.2	6.4	7.45	10.32
1992	75.8	57.8	7.9	7.0	7.41	10.57
1993	75.4	57.9	7.2	6.6	7.39	10.83
1994	75.1	58.8	6.2	6.0	7.40	11.12
1995	75.0	58.9	5.6	5.6	7.40	11.44
1996	74.9	59.3	5.4	5.4	7.43	11.82

资料来源：《总统经济报告》(Economic Report of the President)，1997年

① CLFPRM： 城市劳动力参与率，男性，(%)表B-37。

② CLFPRF： 城市劳动力参与率，女性，(%)表B-37。

③ ONRM：城市失业率，男性，(%)表B-40。

④ ONRM，城市失业率，女性，(%)表B-40。

⑤ AHE82，平均小时工资，(1982美元价)，表B-45。

⑥ AHE，平均小时工资，(当前美元价)，表B-45。

7.15 利用式(7-53)回答下面问题：

R^2值	n	k	$\bar{R}^2$
0.83	50	6	—
0.55	18	9	—
0.33	16	12	—
0.12	1200	32	—

通过上表，你认为R^2与$\bar{R}^2$有什么关系？

7.16 利用式(7-50)建立R^2与$\bar{R}^2$之间的关系。

7.17 对例7.3，计算F值。如果F值是显著的，这意味着什么？

7.18 对例7.2，建立方差分析表并检验假设$R^2=0$，(α=1%)。

7.19 参考习题5.17给出的数据。

(a) 建立一个多元回归模型，解释MBA毕业生的平均初职工资，并且求出回归结果。 (b) 如果模型中包括了GPA和GMAT分数这两个解释变量，先验地，你可能会遇到什么问题，为什么？(c) 如果学费这一变量的系数为正并且在统计上是显著的，是否表示进入最昂贵的商业学校学习是值得的。学费这个变量可用什么来代替？ (d) 假定做GMAT分数对GPA的回归分析，并且发现两变量之间显著正相关。那么，你对多重共线性问题有何看法？ (e) 对(a)建立方差(ANOVA)分析表并检验假设：所有偏回归系数均为零。 (f) 用R^2值，对(e)建立ANOVA表进行分析。

7.20 图7-1表示抵押贷款债务回归结果的正态概率图。

(a) 根据图7-1，你能判定式(7-32)中的误差项服从正态分布吗？为什么？ (b) 观察到的

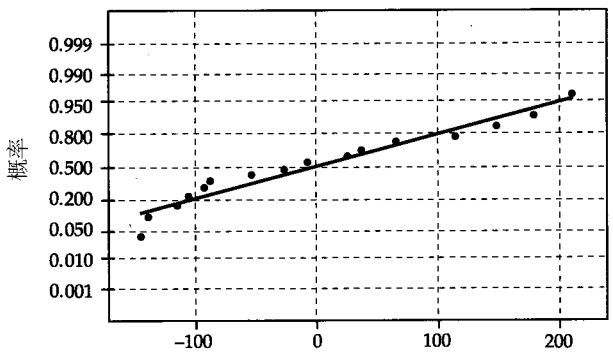


图7-1 正态概率图

Anderson-Darling A^2值0.381在统计上是显著的吗？如果是的，有什么意义？如果不是，你得出什么结论。 (c) 根据表中的数据，你能确定误差项的均值与方差吗？

附录7A.1 OLS估计量的推导

对式(7-16)分别对b_1，b_2求偏导，并令所求偏导为零，得到：

$$\frac{\partial \sum e_i^2}{\partial \sum b_1} = 2\sum (Y_i - b_1 - b_2 X_{2i} - b_3 X_{3i})(-1) = 0$$

$$\frac{\partial \sum e_i^2}{\partial \sum b_2} = 2\sum (Y_i - b_1 - b_2 X_{2i} - b_3 X_{3i})(-X_{2i}) = 0$$

$$\frac{\partial \sum e_i^2}{\partial \sum b_3} = 2\sum (Y_i - b_1 - b_2 X_{2i} - b_3 X_{3i})(-X_{3i}) = 0$$

简化上面这些等式得到式(7-17)，(7-18)和(7-19)。用小写字母表示各值与其均值的离差。(例$x_{2i}=X_{2i}-X_{2i}$)，即可得到式(7-20)(7-21)(7-22)。

附录7A.2 式(7-31)的推导

三变量样本回归模型：

$$Y_i=b_1+b_2X_{2i}+b_3X_{3i}+e_i \tag{7A.2-1}$$

表示成离差形式：

$$y_i=b_1+b_2x_{2i}+b_3x_{3i}+e_i \tag{7A.2-2}$$

(注：$\bar{e}=0$)

因此，

$$e_i=y_i-b_1-b_2x_{2i}-b_3x_{3i} \tag{7A.2-3}$$

则有，

$$\begin{aligned}\sum e_i^2 &= \sum (e_i e_i)\\ &= \sum e_i(y_i - b_1 - b_2 x_{2i} - b_3 x_{3i})\\ &= \sum e_i y_i - b_2 \sum e_i x_{2i} - b_3 \sum e_i x_{3i}\\ &= \sum e_i y_i\\ &= \sum (y_i - b_2 x_{2i} - b_3 x_{3i})(y_i)\\ &= \sum y_i^2 - b_2 \sum y_i x_{2i} - b_3 \sum y_i x_{3i}\\ &= \sum y_i^2 - (b_2 \sum y_i x_{2i} + b_3 \sum y_i x_{3i})\end{aligned} \tag{7-31}$$

附录7A.3 式(7-53)的推导

回顾(参见脚注9)

$$R^2 = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}} \tag{7A.3-1}$$

校正判定系数定义如下：

$$\begin{aligned}\bar{R}^2 &= 1 - \frac{\mathrm{RSS}/(n-k)}{\mathrm{TSS}/(n-1)}\\ &= 1 - \frac{\mathrm{RSS}(n-1)}{\mathrm{TSS}(n-k)}\end{aligned} \tag{7A.3-2}$$

注意其自由度。

现将式(7A.3-1)代入到式(7A.3-2)中，经过代数运算，得

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k}$$

注意：如果不考虑RSS的自由度(=$n-k$)和TSS的自由度(=$n-1$)，虽然，

$$R^2 = \bar{R}^2$$

附录7A.4 EVIEWS计算结果(抵押贷款债务一例)

第 I 部分

LS//Dependent Variable is DEBT
Date: 11/14/97 Time: 09:57
Sample: 1980 1995
Included observations: 16

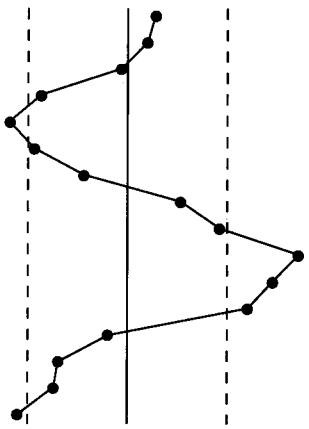
Variable	Coefficient	Std. Error	*t*-Statistic	Prob.	
C	155.6812	578.3288	0.269192	0.7920	(7-32)
Income	0.825816	0.063568	12.99105	0.0000	
Cost	–56.43931	31.4535	–1.794324	0.0960	

R-squared	0.989438	Mean dependent var	2952.175
Adjusted R-squared	0.987813	S.D. dependent var	1132.051
S.E. of regression	124.9716	Akaike info criterion†	9.823534
Sum squared resid	203032.7	Schwarz criterion‡	9.968394
Log likelihood	–98.29128	F-statistic	608.9185
Durbin-Watson stat*	0.402005	Prob (F-statistic)	0.000000

For information on the
*Durbin-Watson Statistic, see Chap. 12.
†Akaike info Criterion, see Chap. 13.
‡Schwarz criterion, see Chap. 13.

第 II 部分

obs	Actual Y_i	Fitted $\hat{Y}_i$	Residual e_i	Residual plot e_i
1980	1365.50	1328.73	36.7719	
1981	1465.50	1440.44	25.0564	
1982	1539.30	1546.34	–7.03706	
1983	1728.20	1834.25	–106.052	
1984	1958.70	2104.12	–145.417	
1985	2228.30	2344.29	–115.985	
1986	2539.90	2593.86	–53.9587	
1987	2897.60	2832.17	65.4309	
1988	3197.30	3082.97	114.329	
1989	3501.70	3291.29	210.412	
1990	3723.40	3545.45	177.952	
1991	3880.90	3732.74	148.165	
1992	4011.10	4037.88	–26.7838	
1993	4185.70	4274.87	–89.1745	
1994	4389.70	4483.95	–94.2549	
1995	4622.00	4761.45	–139.454	



第8章

回归方程的函数形式

到目前为止，我们已经讨论了参数线性模型和变量线性模型。但是，在本书中我们所关注的是参数线性模型，而并不要求变量 Y与X一定是线性的。本章我们将会看到，实际上用以描述许多经济现象的参数线性 /变量线性(LIP/LIV)回归模型可能并不是十分准确的。

举个例子，假定对式 (6-48)给出的LIP/LVP widget需求函数，我们想要估计需求的价格弹性，即，价格每变动 1%所引起的需求量变动的百分比。从式 (6-48)中，我们无法估计出这个弹性，因为模型中的斜率仅仅给出了价格每变动一单位所引起的 (平均的)需求变动的绝对量，但这并不是弹性。然而，对 8.1节所讨论对数线性模型而言，却很容易计算出这样一个弹性。我们会看到，这个模型虽然其参数之间是线性的，但是变量之间却不是线性的。

我们再来看一个例子，假定我们想要求某个经济变量的增长率，[1] 比如国民生产总值(GNP)、货币供给、失业率等。在 8.4节中，你会看到可以用半对数模型 (参数线性，但变量非线性)来度量这个增长率。

需要注意的一点是，即使在参数线性回归模型的限制下，回归模型的形式也有多种。本章将特别讨论下面几种形式的回归模型：

(1) 对数线性模型 (不变弹性模型)。(8.1节)

(2) 半对数模型。 (8.4节、 8.5节)

(3) 双曲函数模型。 (8.6节)

(4) 多项式回归模型。 (8.7节)

所有这些模型的一个重要特征是：它们都是参数线性模型，但变量却不一定是线性的。在第5章中，我们已经讨论过参数之间是线性的模型和变量之间是线性的模型的定义。扼要地说，对于变量之间是线性的模型来说，解释变量每变动一个单位，应变量的变化率为一常数，但是对于变量之间是非线性的回归模型来说，斜率并不是保持不变的。

为了介绍这些基本概念，我们还是从双变量模型着手，然后扩展到多元回归模型。

---

1 如果Y_t和Y_{t-1}是某个变量（比如说 GNP）在t期和$t-1$期的值（比如说， 1997和1996)，那么Y在这两个时期间的增长率为： $\frac{Y_t - Y_{t-1}}{Y_{t-1}}$ 100

它可以简单地理解为 Y的相对或比例变动再乘以 100。在8.4节中将会看到半对数回归模型是如何用于度量在一个较长时期内变量的增长率的。

8.1 如何度量弹性：对数线性模型

我们再来看第 5、第 6章讨论的对 widget需求一例。但是现在考虑如下形式的需求函数。(为使代数形式简洁，随后再引进随机误差项 u_i。)

$$Y_i = AX_i^{B_2} \tag{8-1}$$

在这个模型中，变量 X_i是非线性的。[1]但可将式 (8-1)做恒等变换表示成另一种形式：

$$\ln Y_i = \ln A + B_2 \ln X_i \tag{8-2}$$

其中，ln表示自然对数，即以 e为底的对数。现在若令

$$B_1 = \ln A \tag{8-3}$$

可以将式 (8-2)写为：

$$\ln Y_i = B_1 + B_2 \ln X_i \tag{8-4}$$

为了估计，可将模型 (8-4)写为：

$$\ln Y_i = B_1 + B_2 \ln X_i + u_i \tag{8-5}$$

这是一个线性模型，因为参数 B_1和B_2是以线性形式进入模型的[2]。有趣的是，这个模型还是对数形式变量的线性模型。 (原始模型 (8-1)变量 X是非线性的)，因此，我们将形如式 (8-5)的模型称为双对数(double-log)模型(因为两个变量都以对数形式出现)或对数-线性(log-linear)模型(因为以对数形式出现的变量之间是线性的)。

注意一个“明显的”非线性模型是如何通过适当的变换转变为线性 (参数之间)模型的，这里的变换是对数变换。现令 $Y_i^* = \ln Y_i$, $X_i^* = \ln X_i$ ，则模型 (8-5)可写为：

$$Y_i^* = B_1 + B_2 X_i^* + u_i \tag{8-6}$$

这与我们在前面章节中讨论的模型相似；它不仅是参数线性的，而且变形后的变量 Y^*与X^*之间也是线性的。

对于变形的模型 (8-6)，如果它满足古典线性回归模型的基本假定，则很容易用普通最小二乘法来估计它，并且得到的估计量是最优线性无偏估计量[3]。

在经验工作中，双对数模型 (对数线性模型)应用的非常广泛，其原因在于，它有一个很吸引人的特性：斜率 B_2度量了 Y对X的弹性，即给 X一个(很小)的变动所引起 Y变动的百分比。

如果用符号 ΔY代表 Y的一个小的变动， ΔX代表 X的一个小的变动，定义弹性 E为：

$$E = \frac{Y\text{的变动}\%}{X\text{的变动}\%} = \frac{\Delta Y/Y100}{\Delta X/X100} = \frac{\Delta Y}{\Delta X}\frac{X}{Y}$$

$$= \text{斜率} \times \left(\frac{X}{Y}\right) \tag{8-7}$$[4]

因此，如果 Y代表了商品的需求量， X代表了单位价格，则 E就是需求的价格弹性。

---

1 利用微积分，可以证明： $\frac{dY}{dX} = AB_2X^{(B_2-1)}$。这表明： Y的变化率依赖于 X，并不独立于 X；也就是说， Y的变化率不是一个常数。那么，根据定义，模型（ 8-1）变量 X不是线性的。

2 因为， $B_1 = \ln A$，所以 A可以表示为： $A = \text{antilog}(B_1)$,用数学的语言就是非线性变换。但是，在实际中，截距 A并没有什么经济含义。

3 任何一个统计软件都可对对数进行运算。因此，并不存在额外的计算负担。

4 用积分的形式： $E = \frac{dY}{dX}\frac{X}{Y}$。其中，$dY/dX$表示 Y对X的导数，即 Y对X的变化率。$\Delta Y/\Delta X$是dY/dX的近似。

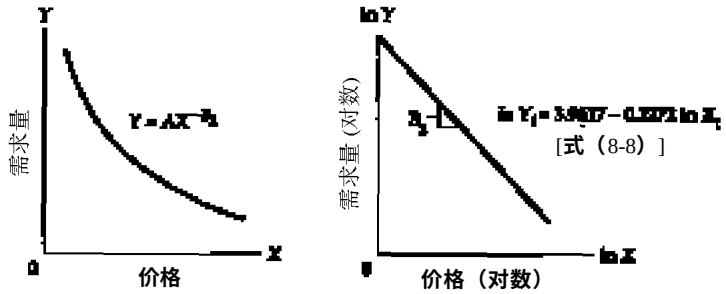
注：对于变形的模型（ 8-6），$B_2 = \frac{\Delta Y^*}{\Delta X^*} = \frac{\Delta \ln Y}{\Delta \ln X} = \frac{\Delta Y/Y}{\Delta X/X} = \frac{\Delta Y}{\Delta X}\frac{X}{Y}$ ，即是 Y对X的弹性。正如本章附录注释，对数形式的改变量就是相对改变量，例如： $\Delta \ln Y = \frac{\Delta Y}{Y}$ 。

我们还可用图形来说明。图 8-1a描绘了函数式 (8-1)，图8-2b是对式(8-1)做对数变形后的图形。图 8-1b中的直线的斜率就是价格弹性的估计值，（$-B_2$)。我们可从图 8-1a中明显地看出对数线性模型的这个重要特征。由于回归线是一条直线（Y和X都是对数形式），所以它的斜率 ($-B_2$)为一常数。对于这个模型，又由于斜率等于其弹性，所以弹性为一常数——它与X的取值无关[1]。

由于这个特殊的性质，双对数模型（对数线性模型）又称为不变弹性模型 (constant elasticity model)。我们将在本书中交替使用这些不同的名称术语。

例8.1　对widget需求。

在(6-48)式中，我们给出了对 widget的需求函数。但是回顾一下散点图，不难发现，需求量和价格之间是近似线性关系的，因为并非所有的样本点都恰好落在直线上。当然，(6-48)式只是为了教学的方便。让我们来看一下，如果用对数线性模型拟合表5-4给出的数据，情况又会怎样？为了方便，这里列出具体的数据，见表 8-1。



a)　　b)

图8-1　不变弹性模型

表8-1　对widget需求

需求量 Y	价格 X	ln$Y$	ln$X$
49	1	3.891 8	0.000 0
45	2	3.806 7	0.693 1
44	3	3.784 2	1.098 6
39	4	3.663 6	1.386 3
38	5	3.637 6	1.609 4
37	6	3.610 9	1.791 8
34	7	3.526 4	1.945 9
33	8	3.496 5	2.079 4
30	9	3.401 2	2.197 2
29	10	3.367 3	2.302 6

OLS回归结果如下：

$$\begin{aligned} \ln Y_i &= 3.9617 - 0.2272 \ln X_i \\ \text{se} &= (0.041\,6) \quad (0.026\,6) \\ t &= (95.233) \quad (-9.0880) \quad r^2 = 0.911\,6 \end{aligned} \tag{8-8}$$

从回归结果可知,价格弹性约为－0.23，表明价格提高1个百分点，平均而言，需求量

1　但是，需谨慎的是，一般，弹性和斜率是两个不同的概念。从式 (8-7)可以清楚地看到，弹性等于斜率乘以 X/Y的比值。仅仅是对于双对数模型（对数线性模型），斜率才等于弹性。

(续)

将下降0.23个百分点。根据定义，如果产品的价格弹性的绝对值小于1，则称该产品是缺乏弹性的。因此，在widget一例中，价格是缺乏弹性的，因为弹性系数的绝对值为0.23。

截距值3.96表示了$\ln X$为零时，$\ln Y$的平均值。同样，这里的截距没有什么具体的经济含义[1]。

r^2=0.916 6，表示$\log X$解释了变量$\log Y$的91%的变动。

图8-1b描绘了式(8-8)所表示的回归直线。

对数线性模型的假设检验

就假设检验而言，线性模型与对数线性模型并没有什么不同。在随机误差项服从正态分布(均值为0，方差为σ^2)的假定下，每一个估计的回归系数均服从正态分布。或者，如果用σ^2的无偏估计量σ^2代替它，则每一个估计的回归系数服从自由度为$(n-k)$的t分布，其中k为包括截距在内的参数的个数。在双变量模型中，k为2，在三变量模型中，k为3，如此等等。

根据式(8-8)的回归结果，读者很容易检验每一个估计的参数在5%的显著水平下，都显著不为零。因为在5%的显著水平下，t值分别为9.08(b_2)，95.26(b_1)，均超过了t临界值2.306(自由度为8，双边检验)。而且获此t值的概率(即p值)非常小。

8.2 线性模型与对数线性模型的比较

下面将要讨论的是一个非常重要的实际问题。我们已经根据样本数据拟合了对widget的需求的线性模型(6-48)以及对数线性模型(8-8)。那么，将选择哪个模型呢？虽然经济理论告诉我们价格与需求量负相关，但是并未提供足够的信息告示这两者之间具体的函数形式。也就是说，经济理论本身并未提供强有力的信息告诉我们是要拟合线性模型、对数线性模型还是其他的模型。那么，回归模型的函数形式就成为一个经验性问题。在选择模型的过程中，是否有规律可循呢？

规律之一是根据数据作图。如果散点图表明两个变量之间的关系近似线性的(也即是一条直线)，那么假定模型是线性的就比较合适。但如果散点图表明变量之间的关系是非线性的，则需要作$\log Y$对$\log X$的图形，如果这个图形表明它们之间是近似线性的，则假定模型是对数线性模型就比较合适。不幸的是，这条规律只适用于双变量情况，对于多变量的情况就不太适合。因为在多维空间中作散点图比较困难。因此，我们需要其他的规则。

为什么不根据r^2来选择模型呢？也就是说，选择r^2值最高的模型。虽说从直观上感觉是可行的，但是这个标准有其自身的问题。首先，在第7章中讲过，要比较两个模型的r^2值，应变量的形式必须是相同的。[2]在线性模型中，r^2度量了X对Y变动解释的比例，但是在对数线性模型中，r^2度量了$\log X$对$\log Y$变动的解释比例。Y的变动与$\log Y$的变动从概念上说是不同的。对数形式度量了一个数的相对变化(或是变化的百分比)，而数字形式度量了该数的绝对变化。[3]因此，对于线性模型(6-48)，X解释了Y变化的98%，但是对于对数线性模型，$\log X$解释了$\log Y$变化的91%。若要比较这两个r^2值，可用习题8.16讨论的方法。

---

1 因为当$\ln X$为零时，$\ln Y$=3.961 7，所以如果对这个数求反对数，为52.546 6。因而，若价格的对数值为零，则对wiget的平均需求量约为53个单位。对线性模型（6-48），这个数字为49.667，约等于50。

2 而与应变量或解释变量的形式无关；它们可以是线性的，也可以不是线性的。

3 若一个数从45到50，则绝对变化为5，但相对变化为（50－45）/45=0.111 1或约为11%。

即使两个模型中的应变量相同，从而两个 r^2 值可直接比较，我们也建议不要根据最高 r^2 值这一标准(high r^2 value criterion)来选择模型。这是因为，正如在第 7章中所指出的那样，r^2(R^2)值可以通过增加进入模型的解释变量的个数而不断增大。研究人员并不是把重点放在模型的 r^2 值上，而是考虑诸如进入模型中的变量之间的相关性 (即以经济理论为基础)、预期的解释变量系数的符号、统计显著性以及类似弹性系数这样的衍生度量工具等因素。这些应该成为选择模型的基本准则。如果按照这些标准，所选模型比其他模型更好，而且若所选模型恰好有较高的 r^2 值，那么就可以认为所选模型是合适的。但是切记要避免仅仅根据 r^2 值的大小来选择模型。

比较式 (8-8)对数线性需求函数与式 (6-48)线性需求函数，不难发现两个模型的斜率均为负值。而且，这两个斜率都是统计显著的。但是，我们不能直接比较这两个斜率，因为在LIV模型中，斜率度量了价格每变动一单位所引起需求量的绝对变化率，而在对数线性模型中，斜率度量的是价格弹性 —— 价格每变动 1%所引起需求量变动的百分比。

如果我们能够计算出 LIV的价格弹性，那么就有可能比较这两个斜率。式 (8-7)表明弹性等于斜率乘以 X与Y的比值 —— 即价格与需求量的比值。虽然，对于线性模型，斜率是一个常数(为什么？)，比如在我们的例子中斜率为－ 2.156 7，但是弹性却因线性需求曲线上的点的不同而变化，因为 X与Y的比值因不同的点而变化。从表 8-1中可以看到，有 10组不同的价格和需求量。因此，原则上可以计算出 10个不同的弹性系数。然而，在实践中，线性模型的弹性系数通常是用平均弹性系数来计算的，即 X和Y的样本均值的比值。

$$\text{平均弹性系数} = \frac{\Delta Y}{\Delta X} \times \frac{\bar{X}}{\bar{Y}} \tag{8-9}$$

其中，$\bar{X}$ 和 $\bar{Y}$ 是X和Y的样本均值。根据表 8-1提供的数据，$\bar{X}$ =5.5，$\bar{Y}$ =37.8。因而，对 widget一例，

$$\text{平均弹性系数} = -2.576 \ \frac{5.5}{37.8} = -0.314$$

有趣的是，对于对数线性需求函数，价格弹性为－ 0.227 2，无论在何种价格水平上来测度价格弹性，价格弹性均保持不变(参见图 8-1b)。这正是为什么称对数线性模型为不变弹性模型的原因。而在另一方面，对于 LIV模型，其弹性系数却随着需求曲线上的点的不同而发生变化[1]。

对线性模型而言，其弹性系数随着需求曲线上的点的不同而变化，而对对数线性模型而言，它在需求曲线上任何一点的弹性系数都是相同的。因此，在这两类模型之间进行选择模型时，我们可以根据这个特点作出判断。因为，在实践中，这些假定都是极端的。很可能在需求曲线上的一小段上的价格弹性是不变的，而在曲线的其他部分上价格弹性却又是个变量。

8.3 多元对数线性回归模型

双变量对数线性回归模型很容易推广到模型中解释变量不止一个的情形。例如，我们可将三变量对数模型表示如下：

$$\ln Y_i = B_1 + B_2 \ln X_{2i} + B_3 \ln X_{3i} + u_i \tag{8-10}$$

在这个模型中，偏斜率系数 B_2、B_3又称为偏弹性系数[2]。因此，B_2是Y对X_2的弹性 (X_3保持不变)，即在X_3为常量时，X_2每变动 1%，Y变化的百分比。由于此时 X_3为常量，所以我们称此弹

1 注意一个有趣的事实：对 LIV模型，其斜率为一常量，而弹性系数却是一个变量。但是，对于对数线性模型，其弹性系数为一常量，而斜率却是一个变量（从脚注 2给出的公式中即可得到）。

2 学过微积分的同学知道，Y对X_2的弹性定义为 $\ln Y$对$\ln X_2$的偏导数：$B_2 = \dfrac{\partial \ln Y}{\partial \ln X_2} = \dfrac{\partial Y / Y}{\partial X_2 / X_2} = \dfrac{\partial Y}{\partial X_2}\dfrac{X_2}{Y}$ 同样地，B_3是Y对X_3的弹性。

性为偏弹性。类似地，B_3是Y对X_3的(偏)弹性(X_2保持不变)。简而言之，在多元对数线性模型中，每一个偏斜率系数度量了在其他变量保持不变的条件下，应变量对某一解释变量的偏弹性。

例8.2 柯布-道格拉斯生产函数

在模型(8-10)中，令Y表示产出，X_2表示劳动投入，X_3表示资本投入。这样，式(8-10)就是一个生产函数——反映产出与劳动力和资本投入之间的关系的函数。这就是著名的**柯布-道格拉斯生产函数**(Cobb-Douglas(C-D)production function)(C-D函数)。我们来看表8-2，它给出了1955～1974年间墨西哥的产出Y，(用国内生产总值GDP度量，以1960年不变价，单位为百万比索)、劳动投入X_2(用总就业人数度量，单位为千人)以及资本投入X_3(用固定资本度量，以1960年不变价，单位为百万比索)的数据。

根据表8-2提供的数据，得到如下回归结果[1]：

$$\begin{aligned}\widehat{\ln Y_t} &= -1.652\ 4 + 0.339\ 7\ln X_{2t} + 0.864\ 0\ln X_{3t} \\ \text{se} &= (0.606\ 2) \quad (0.185\ 7) \quad (0.093\ 4) \\ t &= (-2.73) \quad (1.83) \quad (9.06) \\ p\text{值} &= (0.014)(0.085) \quad (0.000)^{*} \\ R^2 &= 0.995\end{aligned} \tag{8-11}$$

*表示值很小。

**表示F的p值，也很小。

表8-2 实际GDP，就业人数，实际固定资本——墨西哥

年份	GDP①	就业人数②	固定资产③
1955	114 043	8 310	182 113
1956	12 0410	8 529	193 749
1957	129 187	8 738	205 192
1958	134 705	8 952	215 130
1959	139 960	9 171	225 021
1960	150 511	9 569	237 026
1961	157 897	9 527	248 897
1962	165 286	9 662	260 661
1963	178 491	10 334	275 466
1964	199 457	10 981	295 378
1965	212 323	11 746	315 715
1966	226 977	11 521	337 642
1967	241 194	11 540	363 599
1968	260 881	12 066	391 847
1969	277 498	12 297	422 382
1970	296 530	12 955	455 049
1971	306 712	13 338	484 677
1972	329 030	13 738	520 553
1973	354 057	15 924	561 531
1974	374 977	14 154	609 825

资料来源：Victor J.Elias *Sources of Growth:A Study of Seven Latin American Economies*，(International Center for Economic Growth，ICS Press，San Francisco,1992.数据摘自表E5，E12，E14。

① 1960年不变价，单位为百万比索。

② 单位为千人。

③ 1960年不变价，单位为百万比索。

1 该回归结果是用MINTAB统计软件得到的。

(续)

对回归方程(8-11)解释如下。偏斜率系数0.339 7表示产出对劳动投入的弹性。具体说，这个数字表明在资本投入保持不变的条件下，劳动投入每增加一个百分点，平均产出将增加34%。类似地，在劳动投入保持不变的条件下，资本投入每增加一个百分点，产出将平均增加0.85个百分点。如果将两个弹性系数相加，我们将得到一个重要的经济参数——**规模报酬参数**(returns to scale parameter)，它反映了产出对投入的比例变动。如果两个弹性系数之和为 1，则称**规模报酬不变** (constants return to scale)(例如，同时增加劳动和资本为原来的两倍，则产出也是原来的两倍)；如果弹性系数之和大于1，则称**规模报酬递增**(increasing returns to scale)(例如，同时增加两倍的投入，则产出是原产出的两倍多)。如果弹性系数之和小于1，则称**规模报酬递减**(decreasing returns to scale)(例如，同时增加两倍的投入，则产出小于原产出的两倍)。

在本例中，两个弹性系数之和为 1.185 7，表明墨西哥经济的特征是规模报酬递增的。

虽然资本对产出的影响看似大于劳动力对产出的影响，但是根据单边检验的结果，这两个系数各自均是统计显著的。 (注：这里用单边检验，因为我们预期劳动力和资本对产出影响都是正向的)。

估计的F值也是高度相关的(因为p值几乎为零)，因此能够拒绝零假设：劳动力与资本对产出无影响。

R^2值为0.995，表明(对数)劳动力和资本解释了大约 99.5%的(对数)产出的变动。很高的解释程度表明了模型(8-11)很好地拟合了样本数据。

例8.3 对能源的需求

表8-3给出了1960～1982年间7个OECD国家(美国、加拿大、德国、英国、意大利、日本、法国)的总最终能源需求指数(Y)、实际的GDP(X_2)、实际能源价格(X_3)的数据。所有指数均以1970年为基准(1970=100)。利用表8-3提供的数据，得到下面的对数线性需求函数：

$$\begin{aligned}
&\ln \hat{Y}_t = 1.549\ 5 + 0.997\ 2 \ln X_{2t} - 0.331\ 5 \ln X_{3t} \\
&\text{se} = (0.090\ 3) \quad (0.019\ 1) \quad (0.024\ 3) \\
&t = (17.17) \quad (52.09) \quad (13.61) \\
&p\text{值} = (0.000)^* \quad (0.000)^* \quad (0.000)^* \qquad (8\text{-}12) \\
&R^2 = 0.994 \\
&\bar{R}^2 = 0.994 \\
&F = 1688
\end{aligned}$$

*表示值很小。

回归结果表明，能源的需求与收入 (用实际GDP度量)正相关，与实际价格负相关；这与经济理论相符。收入弹性的估计值为 0.99，表示在其他条件保持不变的条件下，实际收入每增加一个百分点，对能源的平均需求量将增加 0.99%。同样地，在其他因素保持不变的条件下，能源价格每上涨 1%，则对能源的平均需求量将降低 0.33

(续)

个百分点。由于这个系数的绝对值小于 1，因此，可以认为能源的需求对价格是缺乏弹性的，这并没有什么可奇怪的，因为能源是最基本的消费品。

校正的和未校正的R^2值都很高。F值为1 688，也很高；若$B_2=B_3=0$为真，则获此F值的概率几乎为零。因此，我们能够认为收入和能源价格对能源的需求有很强的影响。

表8-3 OCED国家对能源的需求

年份	最终需求	实际的GDP	实际的能源价格
1960	54.1	54.1	111.9
1961	55.4	56.4	112.4
1962	58.5	59.4	111.1
1963	61.7	62.1	110.2
1964	63.6	65.9	109.0
1965	66.8	69.5	108.3
1966	70.3	73.2	105.3
1967	73.5	75.7	105.4
1968	78.3	79.9	104.3
1969	83.3	83.8	101.7
1970	88.9	86.2	97.7
1971	91.8	89.8	100.3
1972	97.2	94.3	98.6
1973	100.0	100.0	100.0
1974	97.3	101.4	120.1
1975	93.5	100.5	131.0
1976	99.1	105.3	129.6
1977	100.9	109.9	137.7
1978	103.9	114.4	133.7
1979	106.9	118.3	144.5
1980	101.2	119.6	179.0
1981	98.1	121.1	189.4
1982	95.6	120.6	190.9

资料来源：Richard D.Prosser，“Demand Elasticities in OECD:Dynamic Aspects，”*Energy Economics*，January 1985.p.10.

8.4 如何测度增长率：半对数模型

在本章前言中曾提到，通常经济学家、工商业家和政府对某一经济变量的增长率很感兴趣。比如说，政府预算赤字规划就是根据预计的 GNP增长率这一最重要的经济活动指标而确定的。类似地，联储根据未偿付消费者信贷的增长率 (自动贷款、分期偿还贷款等等) 这一指标来监视其货币政策的运行效果。

本节将介绍回归分析是如何用于测度这些增长率的。

例8.4 1973～1987年间美国未偿付消费者信贷的增长

表8-4给出了美国1973～1987年间未偿付消费者信贷的数据。

我们现在要求在此期间的未偿付消费者信贷的增长率（Y）。我们来看货币、银行及金融等课程中介绍过的复利计算公式：

$$Y_t = Y_0 (1+r)^t \tag{8-13}$$[1]

其中，Y_0 —— Y的初始值

Y_t —— 第t期的Y值

r —— Y的增长率（复利率）

将式(8-13)变形，对等式两边取对数，得：

$$\ln Y_t = \ln Y_0 + t\ln(1+r) \tag{8-14}$$

现令

$$B_1 = \ln Y_0 \tag{8-15}$$

$$B_2 = \ln(1+r) \tag{8-16}$$

因此，模型(8-14)可表示为：

$$\ln Y_t = B_1 + B_2 t \tag{8-17}$$

若引进随机误差项，得到：[2]

$$\ln Y_t = B_1 + B_2 t + u_t \tag{8-18}$$

形如式(8-18)的回归模型称为半对数模型，因为仅有一个变量以对数形式出现（在这个例子中，应变量以对数形式出现）。如何解释这类模型呢？在解释半对数模型之前，注意，在满足OLS基本假定的条件下，能够用普通最小二乘法来估计模型（8-18）。根据表8-4提供的数据，得到如下回归结果：

$$\widehat{\ln Y_t} = 12.007 + 0.094\,6t \tag{8-19}$$

$$\text{se} = (0.031\,9) \qquad (0.003\,5)$$

$$t = (376.40) \qquad (26.03) \qquad r^2 = 0.982\,4$$

图8-2a描绘了估计的回归直线。

表8-4 美国未偿付消费者信贷（Y）*

年份	Y	年份	Y
1973	190 601	1981	366 597
1974	199 365	1982	381 115
1975	204 963	1983	430 382
1976	228 162	1984	511 768
1977	263 808	1985	592 409
1978	308 272	1986	646 055
1979	347 507	1987	685 545
1980	349 386		

注：*单位为百万美元。

资料来源：《总统经济报告》，1989年，摘自表B-75，第396页。

1 假定在你的存折上的存款 Y_0=100美元，假设年利率为 6%，即r=6%。在第1年年末，银行存款将增加到 Y_1=100（1+0.06）=106；在第2年年末，存款为Y_2=106（1+0.06）=100（1+0.06）^{2}=112.36，因为在第2年，你不但得到最初的100美元的利息，而且还有第1年所得的利息。第3年，存款为100（1+0.06）^{3}=119.101 6，等等。

2 加上误差项是因为复利率的计算公式并未完全拟合表 8-4的数据。

(续)

对式(8-19)回归结果解释如下。斜率0.094 6表示，平均而言，log*Y*(未偿付消费者信贷)的年增长率为0.094 6。用一般的语言解释，即为*Y*的年增长率为9.46%，因为，在诸如式(8-19)这样的半对数模型中，斜率度量了给定解释变量的绝对变化所引起的*Y*的比例变动或相对变动[1]。将此相对改变量乘以100，就得到增长率(参见脚注1)。在这个例子中，相对变化率为0.094 6，因而增长率为9.46%。

正因如此，所以半对数模型又称为增长模型，通常我们用这类模型来测度许多变量的增长率，包括经济变量和其他一些非经济变量。

对截距12.007解释如下。根据式(8-15)，显然有：

$$b_1=\ln Y_0\text{的估计值}=12.007$$

因此，若取12.007的反对数，得：

$$\text{antilog}\ (12.007)\approx 163\ 911.7 \tag{8-20}$$

即是当$t=0$时的Y值，也即Y的初期值。从表8-4可知，1973年底，即初期值为190 000(百万美元)。因而，我们或许能将截距值163 912(百万美元)解释为1973年初始时期的未偿付消费者信贷量。但别忘了在前面，我们多次指出：通常截距没有特别实际的意义。

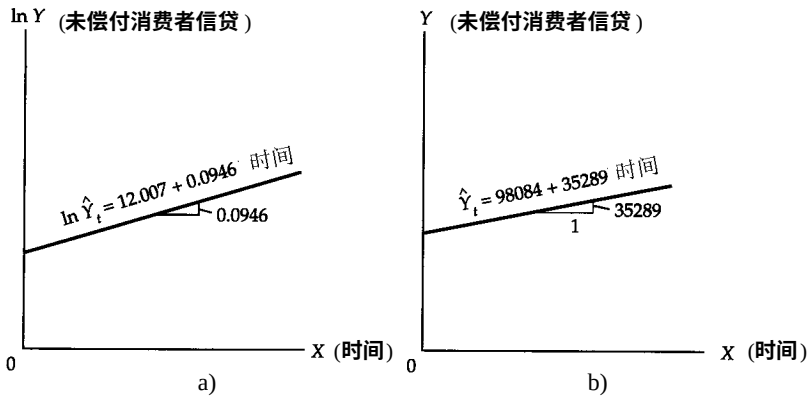


图8-2 未偿付消费者信贷的增长率
a)半对数模型 b)线性模型

8.4.1 单利增长率与复利增长率

注意观察式(8-16)：

$$b_2=B_2\text{的估计值}=\ln(1+r)$$

因此，

$$\text{antilog}(b_2)=(1+r)$$

于是，

$$r=1-\text{antilog}(b_2) \tag{8-21}$$

由于r是复利增长率，因此一旦计算出b_2值，就很容易根据式(8-21)估计出Y的复利增长率。在这个例子中，

1 利用微分，可以证明：$B_2=\dfrac{d\ln Y}{dt}=\left(\dfrac{1}{Y}\right)\left(\dfrac{dY}{dt}\right)=\dfrac{\dfrac{dY}{dt}}{Y}=\dfrac{Y\text{的相对变化}}{Y\text{的绝对变化}}$

$$r=\text{antilog}\ (0.094\ 6)-1=1.0922-1=0.099\ 2 \tag{8-22}$$

即在样本区间内，未偿付消费者信贷的年增长率为 9.92%。

在前面求得 Y的增长率为 9.46%，但现在却为 9.92%。这有什么区别呢？ 9.46%是单利增长率[或者，更一般地，将形如回归式 (8-19)的斜率系数乘以 100]，而9.92%是复利(一段时期内)增长率。

然而，在实际中，我们通常列出的是单利增长率(instantaneous growth rate)，虽然复利增长率(compound growth rate)很容易计算。

8.4.2 线性趋势模型

有时为了计算的简便，研究人员对下面的模型进行估计：

$$Y_t=B_1+B_{2t}+u_t \tag{8-23}$$

即Y对时间 t的回归，其中 t按时间先后顺序计算。这类模型称为线性趋势模型，时间 t称为趋势变量[1]。若上式中的斜率为正，则称 Y有向上的趋势，若斜率为负，则称 Y有向下的趋势。

根据表8-4提供的数据，拟合的回归方程如下：

$$\begin{aligned}\hat{Y}_i &= 98\ 084+35\ 289t \\ \text{se} &= (23\ 095) \quad (2\ 540.1) \\ t &= (4.247\ 0) \quad (13.893) \quad r^2=0.936\ 9 \\ p\text{值} &= (0.000)^* \quad (0.000)^*\end{aligned} \tag{8-24}$$

*表示值很小。

回归结果表明，在样本区间内，未偿付消费者信贷的年绝对增长率为 35 289(百万美元)。因此，在此期间，未偿付消费者信贷有一个向上的趋势。 (参见图8-2b)。

在实际中，线性趋势模型和增长模型都应用的很广泛。但相比较而言，增长模型更有用一些。人们通常关注的是经济变量的相对变化而不是绝对变化，比如说， GNP，货币供给等等。

同时要注意的是不能比较两个模型的 r^2值，因为两个模型的应变量不相同 (参见习题8.16)。用统计的语言来说，根据 t显著性检验，两个模型的回归结果都相当好。

回顾对数线性模型 (或双对数模型)，斜率也即是 Y对相关解释变量的弹性系数。对增长模型和线性趋势模型，我们也可以测度其弹性系数。事实上，一旦知道回归模型的函数形式，就能根据式 (8-7)给出的弹性的基本定义来计算弹性系数。本章最后的表 8-8总结了本章介绍过的各种模型的弹性系数。

注意：近来，新一代时间序列经济计量学家对诸如模型 (8-18)、(8-23)中的趋势变量法提出了疑问。他们认为这样引进时间变量 t，只有当随机项 u_t在上述模型中是固定不变时才能成立。我们将在随后的章节中解释固定不变这一概念的准确的含义，但是，如果 u_t的均值和方差不随时间的变化而变化，就可以认为 u_t是固定不变的。在古典线性回归模型中，已经假设了 u_t的均值为零，方差为一常量 σ^2。当然，在实际应用中，我们要检查这些假定是否是有效的。我们将在本书后面的章节讨论这个问题。

8.5 线性对数模型：解释变量是对数形式

在上一节，我们讨论了应变量是对数形式而解释变量是线性形式的增长模型。为了描

1 趋势描述了变量持续向上或向下运动的行为。

述的方便，称之为对数-线性模型(log-lin model)或增长模型(growth model)。本节我们将考虑应变量是线性形式而解释变量是对数形式的模型。相应地，我们称之为线性-对数模型(lin-log model)。

我们用一个具体例子来介绍这个模型。

例8.5 美国GNP与货币供给间的关系(1973~1987年)

假定联储很关注货币供给的变动对GNP的影响(货币供给是由FED控制的)。表8-5给出了GNP和货币供给(用M_2度量)的数据。

现考虑下面模型：

$$Y_t=B_1+B_2\ln X_{2t}+u_t \tag{8-25}$$

其中，Y=GNP，X=货币供给。

与对数线性模型相比，对数线性模型中的应变量是对数形式，解释变量是线性形式。在解释线性对数模型之前，先给出模型(8-25)的回归结果：

$$\hat{Y}_t=-16\ 329.0+2\ 584.8\ln X_t \tag{8-26}$$

$$t=(-23.494)\qquad(27.549)\qquad r^2=0.983\ 2$$

按照通常的解释，斜率系数2 585表示货币供给每增加一个百分点，GNP的绝对变化量为2587亿美元。用日常的语言如何解释呢？

回顾一下：对数形式的变化称为相对变化。因此，模型(8-25)中的斜率系数度量了[1]：

$$B_2=\frac{Y\text{的绝对变化量}}{X\text{的相对变化量}}=\frac{\Delta Y}{\Delta X/X} \tag{8-27}$$

与前面相同，其中，ΔY和ΔX表示了Y和X的一个(小的)改变量。式(8-27)也可以写为：

$$Y=B_2\left(\frac{\Delta X}{X}\right) \tag{8-28}$$

式(8-28)表明，Y的绝对变化量等于B_2乘以X的相对变化量。若将后者乘以100，则式(8-28)给出了X每变动一个百分点，Y的绝对变动量。因而，若$\Delta X/X$每变化0.01个单位(或1%)，则Y的绝对变化量为0.01(B_2)。若在实际中，求得B_2=674，则Y的绝对改变量为0.01(674)=6.74。因此，在用OLS方法估计形如式(8-25)的回归方程时，需将估计的斜率系数B_2乘以0.01，或除以100。

我们再来看模型(8-26)给出的GNP/货币供给的回归结果，不难发现货币供给每增加一个百分点，平均而言，GNP将增加25.84亿美元(注：将估计的斜率值除以100)。

因而，形如式(8-25)的线性对数模型常用于研究解释变量每变动1%，相应应变量的绝对变化量的情形。无须赘言，形如(8-25)的模型可以有不止一个的对数形式的解释变量。每一个偏斜率系数度量了在其他变量保持不变的条件下，某一给定变量X每变动1%所引起的应变量的绝对改变量。

---

1 如果$Y=B_1+B_2\ln X$，用微分，可以证明：$\frac{dY}{dX}=B_2\left(\frac{1}{X}\right)$

因此，$B_2=X\frac{dY}{dX}=\frac{dY}{dX/X}$ =式(8-27)

8.6　双曲函数模型

形如下式的模型称为双曲函数模型：

$$Y_i = B_1 + B_2\left(\frac{1}{X_i}\right) + u_i \tag{8-29}$$

这是一个变量之间是非线性的模型，因为 X是以倒数形式进入模型的，但这个模型却是参数线性模型，因为模型中参数之间是线性的[1]。

这个模型的一个显著特征是，随着 X的无限增大，$(1/X_i)$将接近于零 (为什么？)，Y将逐渐接近B_1渐进值(asymptotic value)或极值。因此，当变量 X无限增大时，形如式 (8-29)的回归模型将逐渐靠近其渐进线或极值。

图8-3给出了双曲函数模型的一些可能的形状。

表8-5　美国GNP与货币供给

年份	GNP/亿美元	M_2
1973	1 359.3	861.0
1974	1 472.8	908.5
1975	1 598.4	1 023.2
1976	1 782.8	1 163.7
1977	1 990.5	1 286.7
1978	2 249.7	1 389.0
1979	2 508.2	1 500.2
1980	2 723.0	1 633.1
1981	3 052.6	1 795.5
1982	3 166.0	1 954.0
1983	3 405.7	2 185.2
1984	3 772.2	2 363.6
1985	4 014.9	2 562.6
1986	4 240.3	2 807.7
1987	4 526.7	2 901.0

资料来源：《总统经济报告》，1989，GNP数据摘自表 B-1，第308页，M_2数据摘自表 B-67，第385页。

注：GNP数据根据年利率进行了季节调整

M_2=通货 +活期存款 +旅游支票 +其他可核对存款 +短期 RPs与欧元 +MMMF余额 +MMDAs+储蓄与小额存款

这些都是平均每天的数据，经过季节调整。

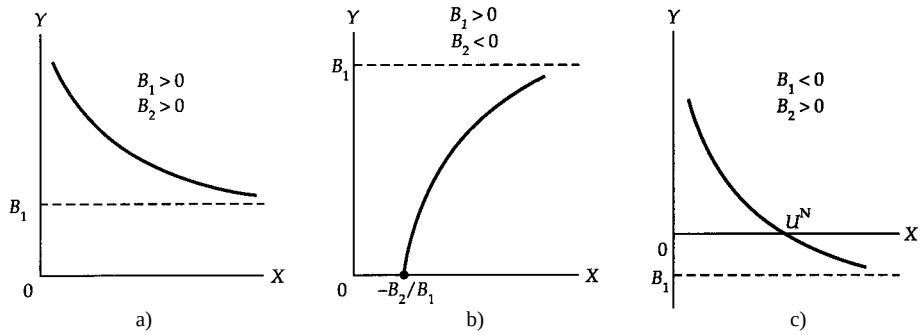


图8-3　双曲函数模型：$Y_i=B_1+B_2(1/X_i)$

1　如果令 $X_i^* = (1/X_i)$，则式（8-29）不仅参数之间是线性的，而且变量之间也是线性的。

在图8-3a中，若Y表示生产的平均固定成本（AFC），也即总固定成本除以产出，X代表产出，则根据经济理论，随着产出的不断增加，AFC将逐渐降低（因为总固定成本不变），最终接近其渐进线（$X=B_1$）。

图8-3b的一个重要用途是它描绘了恩格尔消费曲线（Engel expenditure curve）（以德国统计学家恩格尔的名字命名，1821~1896）。该曲线表明：消费者对某一商品的支出占其总收入或总消费支出的比例。若Y表示消费者在某一商品上的消费支出，X表示消费者总收入，则该商品有如下特征：（1）收入有一个临界值或阈值，在此临界值之下，不能购买某商品（比如汽车）。在图8-3b中，收入的临界水平是-（B_2/B_1）。（2）消费有一个满足水平，在此水平之上，无论消费者的收入有多高，也不会有任何消费（即使是百万富翁通常也不会在同一时间内拥有两辆以上的汽车）。在图8-3b中，消费的满足水平为渐进线$X=B_1$。双曲函数是描述这类商品最合适的模型。

图8-3c的一个重要用途是宏观经济学中著名的菲利普斯曲线（Philips curve），菲利普斯根据英国货币工资变化的百分比（Y）与失业率（X）的数据，得到了形如图8-3c的一条曲线[1]。从图中可以看出，工资的变化对失业水平的反映是不对称的：失业率每变化一个单位，则在失业率低于自然失业率U^N水平时的工资上升的比在当失业率在自然失业率水平以上时快。B_1表明了渐进线的位置。菲利普斯曲线这条特殊的性质可能是由于制度的因素，比如工会交易势力、最少工资、失业保险等等。

例8.6 1958~1969年美国的菲利普斯曲线

由于菲利普斯曲线的历史重要性，也为了阐明双曲函数模型，我们来看一个具体实例。表8-6给出了美国1958~1969年间小时收入指数（Y）和城市失业率（X）的数据。

模型（8-29）拟合了表8-6给出的数据，回归结果如下：

$$\hat{Y}_t = -0.259\ 4 + 20.588\ 0\left(\frac{1}{X_t}\right) \qquad (8\text{-}30)$$

$$t=(-0.257\ 2) \qquad (4.399\ 6) \qquad r^2=0.659\ 4$$

图8-4a给出了该回归线。

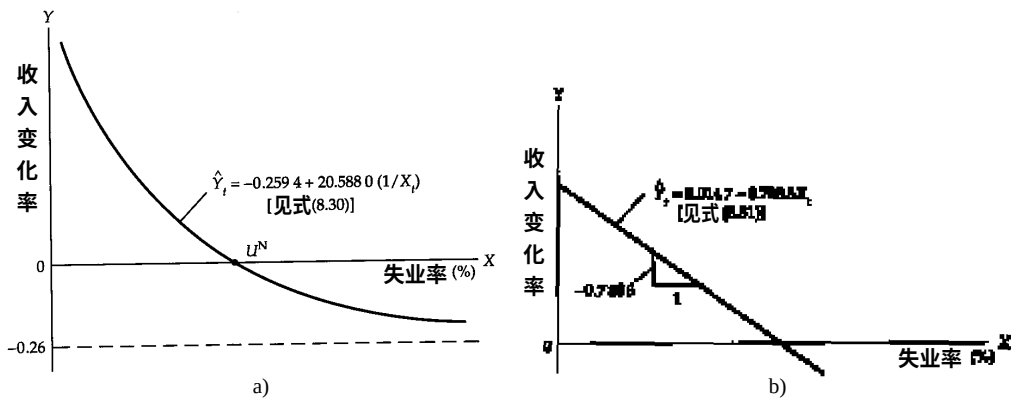


图8-4 1958~1969年美国的菲利普斯曲线

1 菲利普斯（A.W.Phillips），《英国失业与货币工资变动率之间的关系（1861~1957）》（The Relation between Unemployment and the Rate of Money Wages in the United Kingdom，1861~1957）《经济学家》（*Economica*），1958年9月，第283-299页。

(续)

从图中可以看出，最低工资为－ 0.26%，它显著不为零(为什么？)。因此，无论失业率有多高，工资的增长率至多为零。

作为比较，我们给出根据相同数据得到的线性的回归结果：

$$\hat{Y}_t = 8.014\ 7 - 0.788\ 3X_t \tag{8-31}$$

$$t=(6.462\ 5)\quad(-3.260\ 5)\qquad r^2=0.515\ 3$$

观察这两个模型。在线性模型(8-31)中，斜率为负，因为在其他条件保持不变时，失业率越高，收入的增长率越低。然而，在双曲函数模型中，斜率却为正，这是因为X以倒数形式进入模型的(也就是说，负负为正)。换句话说，在双曲函数模型中的正的斜率与在线性模型中负的斜率的作用相同。线性模型表明，失业率每上升 1%，平均而言，无论如何度量X，收入的变化率为一常数，约为-0.79。而另一方面，在双曲函数模型中，收入的变化率却不是常数，它依赖于 X(即失业率)的水平(参见表8-8)[1]。后一种模型看似更符合经济理论。由于在两个模型中的应变量相同，因此我们可以比较r^2值。双曲函数模型的r^2值比线性模型的大，这表明前者比后者更好地拟合了样本数据。

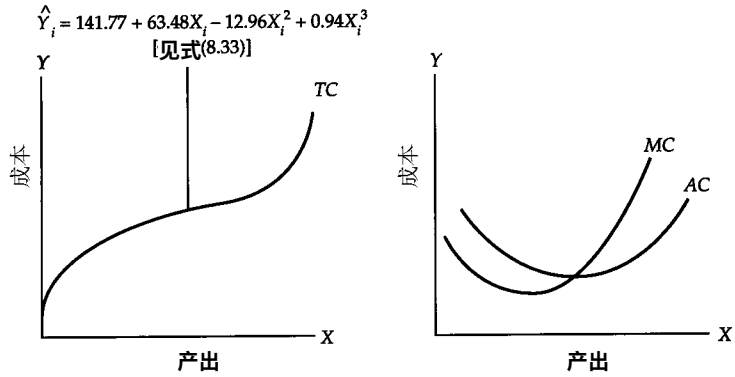


图8-5　成本-产出关系

这个例子表明，一旦遇到的不是 LIV/LIP模型，而是参数线性而变量却不一定是线性的模型，我们需要根据具体情况来仔细地选择合适的模型。在选择时，现象背后的经济理论对选择适当的模型有很大的帮助。不可否认，建立模型需要经济理论和实际经验。而后者来源于不断的实践。

8.7　多项式回归模型

本节讨论的这类回归模型在生产与成本函数这个领域中被广泛地使用。图 8-5描绘了总成本函数(是产出的函数)曲线和边际成本(*MC*)及平均成本(*AC*)曲线。

*Y*表示总成本(*TC*)，*X*表示产出，总成本函数可表示为：

$$Y_i = B_1 + B_2X_i + B_3X_i^2 + B_4X_i^3 \tag{8-32}$$

1　见表8-8，双曲函数模型的斜率为－ $B_2(1/X_i^2)$，显然不是常数。

形如式(8-32)的函数又称为立方函数(cubic function)，一般也称，三次多项式函数(third-degree polynomial)——变量X的最高次幂代表了多项式函数的次（此时的最高次为3)。

注意：在这类多项式函数中，等式右边只有一个解释变量，但却以不同的次幂出现，因而，可把它们看做多元回归模型[1]。(引入随机误差项，则模型(8-32)就成为回归模型。)

虽然，模型(8-32)的变量之间是非线性的，但其参数B却是线性的，因此它是线性回归模型。所以我们可用普通最小二乘法来估计形如(8-32)的模型。惟一的担心是：可能出现自相关问题，因为X的不同次方项是函数相关的。但这种担心没什么必要，因为X^2和X^3是X的非线性函数，因而并未违背不完全共线性的假定，也即变量之间不完全共线性。简言之，多项式回归模型可以用普通最小二乘法估计并且不会带来任何特别的估计问题。

表8-6 美国小时收入指数年变化的百分比(Y)与失业率(%)(X)

年份	Y	X	年份	Y	X
1958	4.2	6.8	1964	2.8	5.2
1959	3.5	5.5	1965	3.6	4.5
1960	3.4	5.5	1966	4.3	3.8
1961	3.0	6.7	1967	5.0	3.8
1962	3.4	5.5	1968	6.1	3.6
1963	2.8	5.7	1969	6.7	3.5

资料来源：《总统经济报告》，1989，Y的数据摘自表B-39，第352页，X的数据摘自表B-44，第358页。

例8.7 总成本函数：为了说明多项式模型，考虑表8-7给出的成本-产出数据

根据这些数据，用普通最小二乘法得到的回归结果如下：

$$\hat{Y}_i = 141.7667 + 63.4776\,X_i - 12.9615\,X_i^2 + 0.939\,6\,X_i^3$$

$$\text{se}=(6.375\ 3) \quad (4.778\ 6) \quad (0.985\ 7) \quad (0.059\ 1)$$

$$R^2=0.998\ 3 \qquad (8\text{-}33)$$

如果边际成本的曲线和平均成本的曲线为U型，根据价格理论可知，模型中的系数有如下先验值[2]：

1.B_1，B_2和B_4都大于零。

2.$B_3<0$

3.$B_3^2<3B_2B_4$

式(8-33)回归结果与这些预期一致。

表8-7 成本-产出数据

Y($)	193	226	240	244	257	260	274	297	350	420	总成本
X	1	2	3	4	5	6	7	8	9	10	产出

8.8 不同函数形式模型小结

本章我们讨论了几种回归模型，这些模型的参数之间是线性的，但变量之间却不一定是线性的。对每一种模型，我们指出了其特殊的性质并强调了其适用条件。表8-8总结了

1 当然，如果需要的话，可以将模型中引入其他变量以及它的不同次幂。

2 参见Alpha C.Chiang的《经济数学的基本方法》(Fundamental Methods of Mathematical Economics)。第三版，McGraw-Hill,纽约，1984，第205～252页。这些限制条件是为了保证有经济意义，即总成本曲线必须向上倾斜（投入越大，产出越高），边际成本必须为正。

讨论过的不同函数形式模型的一些显著特征，比如斜率系数和弹性系数。虽然，对双对数模型而言，其斜率系数与弹性系数相同，但对于其他模型，情况却不如此。但对每一种模型，我们可以根据弹性系数的基本定义 [见式(8-7)]来计算其弹性系数的大小。

表8-8 不同函数形式模型小结

模 型	形 式	斜率 $=\frac{dY}{dX}$	弹性 $=\frac{dY}{dX}\frac{X}{Y}$
线性模型	$Y=B_1+B_2X$	B_2	$B_2\left(\frac{X}{Y}\right)^*$
双对数模型	$\ln Y=B_1+B_2\ln X$	$B_2\left(\frac{Y}{X}\right)$	B_2
对数-线性模型	$\ln Y=B_1+B_2X$	B_2Y	$B_2(X)^*$
线性-对数模型	$Y=B_1+B_2\ln X$	$B_2\left(\frac{1}{X}\right)$	$B_2\left(\frac{1}{X}\right)^*$
双曲函数模型	$Y=B_1+B_2\left(\frac{1}{X}\right)$	$-B_2\left(\frac{1}{X^2}\right)$	$-B_2\left(\frac{1}{XY}\right)^*$

注：*表示弹性系数是一个变量，其值依赖于 X或Y或X与Y。在实际中，若没有给出具体的 X，Y，则用 $\bar{X},\bar{Y}$ 来测度弹性系数。

从表8-8可以看出，对变量之间是线性的模型，其斜率为一常数，而弹性系数是一个变量，但对双对数模型，弹性系数是一常数，而斜率为一变量。表 8-8中的其他模型，斜率和弹性系数都是变量。

8.9 小结

本章我们讨论了几种参数之间是线性的模型或是可通过适当的变形成参数线性的模型(但变量之间并不一定是线性的)。这些不同的模型都有其自己特殊的性质和特征。主要的 5 种模型如下：

(1) 双对数模型。应变量和解释变量都是对数形式。

(2) 对数-线性模型。应变量是对数形式，但解释变量是线性形式。

(3) 线性-对数模型。解释变量是对数形式，但应变量是线性形式。

(4) 双曲函数模型。应变量是线性形式，但解释变量是倒数形式 [1]。

(5) 多项式模型。解释变量以不同次方形式进入模型。

当然，研究人员可以将上述不同形式的模型联合起来。因而，就可以得到多元回归模型，即应变量是对数形式，有些解释变量是对数形式，有些解释变量是线性形式。

我们研究了这些不同模型的性质，并给出了不同模型的适用条件。无须赘言，我们将在本书后面章节中遇到这些模型。

在选择模型时，不应过分强调而且不要仅仅根据一个统计量，比如 R^2来判定。模型的建立需要正确的理论，合适有用的数据，对各种不同模型统计性质的完整的理解以及经验判断。由于理论本身不非完美的，因此也就没有完美的模型，我们只期望选择模型时能够合理地权衡上述各项标准。

习题

8.1 简要解释下列概念

(a) 双对数模型 (b) 对数-线性模型 (c) 线性-对数模型 (d) 弹性系数 (e) 均值弹性

1 也可以是应变量是倒数形式，而解释变量是线性形式，见习题 8.15和8.20。

8.2 什么是斜率系数和弹性系数？两者之间有什么联系？

8.3 填充下表：

模型	适用条件
$\ln Y_i = B_1 + B_2 \ln X_i$	-
$\ln Y_i = B_1 + B_2 X_i$	-
$Y_i = B_1 + B_2 \ln X_i$	-
$Y_i = B_1 + B_2 \left(\frac{1}{X_i}\right)$	-

8.4 完成下列各句：

(a) 在双对数模型中，斜率测度了…… (b) 在线性-对数模型中，斜率测度了…… (c) 在对数-线性模型中，斜率测度了…… (d) Y对X的弹性定义为…… (e) 价格弹性的定义为…… (f) 需求称为富有弹性的，如果价格弹性的绝对值……，需求称为缺乏弹性的，如果价格弹性的绝对值……

8.5 判断正误并说明理由。

(a) 双对数模型的斜率和弹性系数相同。 (b) 对于变量之间是线性的模型 (LIV)而言，斜率系数是一个常数，弹性系数是一个变量。但双对数模型的弹性系数是一个常数，而斜率是一个变量。 (c) 双对数模型的 R^2值可以与对数-线性模型的相比较，但不能与线性-对数模型的相比较。 (d) 线性-对数模型的 R^2值可以与线性模型相比较，但不能与双对数模型或对数线性模型的相比较。 (e) 模型 A：$\ln \hat{Y} = -0.6+0.4X$；$r^2=0.85$；模型 B：$\hat{Y} =1.3+2.2X$；$r^2=0.73$模型A更好一些，因为它的 r^2大。

8.6 恩格尔曲线表明了一个消费者对某一商品的消费支出占总收入的比重。令 Y表示在某一商品上的消费支出， X表示消费者收入，考虑下面的模型：

(a) $Y_i=B_1+B_2X_i+u_i$ (b) $Y_i=B_1+B_2(1/X_i)+u_i$ (c) $\ln Y_i=B_1+B_2\ln X_i+u_i$ (d) $\ln Y_i=B_1+B_2(1/X_i)+u_i$ (e) $Y_i=B_1+B_2\ln X_i+u_i$

你将选择哪个模型？ (提示：解释各个斜率，求出各个支出对收入的弹性系数的表达式)

8.7 根据几个经济时间序列数据 (US)，增长模型 (8-18)的回归结果如下：

时间序列	B_1	B_2	r^2
实际GNP(1954~1958)	7.249 2	0.030 2	0.983 9
(1982美元)	t=(529.29)	(44.318)	
劳动力参与率	4.105 6	0.053	0.946 4
(1973~1987)	t=(1 290.8)	(15.149)	
S&P500指数	3.696 0	0.045 6	0.863 3
(1954~1987)	t=(57.408)	(14.219)	
S&P500指数	3.711 5	0.011 4	0.852 4
(1954~1987，季数据)	t=(114.615)	(27.819)	

(a) 求各单利增长率。 (b) 求各复利增长率 (c) 对S&P数据，为什么两个斜率不相同？如何协调这个差距？

8.8 参考式(8-33)给出的三次方形式的总成本函数 (TC)，

(a) 边际成本函数 (MC)是指产出每变化一单位，总成本的改变量；也即，总成本对产出的变化率。 (专业地，它是总成本对产出的导数。)根据式 (8-33)推倒出边际成本函数。 (b) 平均可变成本 (AVC)是总可变成本 (TVC)除以总产出。根据式 (8-33)推倒出平均可变成本

函数。 (c) 平均成本 (AC)是总成本除以总产出。根据式 (8-33)推倒出平均成本函数。(d) 作出上述各种成本曲线，并验证这些成本曲线与课本上的标准成本曲线相类似。

8.9 下面的模型是参数之间线性的吗？如果不是，有什么方法使之成为参数线性模型？

(a) $Y_i = \dfrac{1}{B_1 + B_2 X_i}$ (b) $Y_i = \dfrac{X_i}{B_1 + B_2 X_i^2}$

8.10 根据11个年观察值，得到下面的回归模型：

模型A： $\hat{Y}_t = 2.691\ 1 - 0.479\ 5 X_t$

se=(0.121 6) (0.114 0) r^2=0.662 8

模型B： $\ln \hat{Y}_t = 0.777\ 4 - 0.253\ 0 \ln X_t$

se=(0.015 2) (0.049 4) r^2=0.744 8

其中，Y表示每人每天消费咖啡的杯数， X表示咖啡的价格 (美元/磅)。

(a) 解释这两个模型的斜率系数。 (b) 已知 $\bar{Y} = 2.43$, $\bar{X} = 1.11$。根据这些值估计模型A的价格弹性。 (c) 求模型B的价格弹性？ (d) 从估计的弹性看，你是否能说咖啡的需求对价格是缺乏弹性的？ (e) 如何解释模型 B的截距？(提示：取反对数) (f) “由于模型B的r^2值比模型A的大，所以模型B比A好。”这句话对吗？为什么？

8.11 参考式(8-11)给出的C-D生产函数：

(a) 解释劳动投入 (X_2)的系数。它显著不为1吗？ (b) 解释资本投入 (X_1)的系数。它显著不为0吗？显著不为1吗？ (c) 截距-3.338 5有什么意义？ (d) 检验假设： $B_2 = B_3 = 0$。

8.12 Mohsen Bahami-Oskooee和Margaret Malixi [1]在研究28个不发达国家(LDCs)对国际储备(即外汇储备，例如美元或国际货币基金组织的特别提款权)需求时，得到下面的回归结果：

$$\widehat{\ln(R/P)} = 0.122\ 3 + 0.407\ 9 \ln(Y/P) + 0.504\ 0 \ln \sigma_{BP} - 0.091\ 8 \ln \sigma_{EX}$$

t = (2.512 8) (17.637 7) (15.243 7) (-2.749 9)

R^2=0.826 8

F=1 151

n=1 120

其中，

R—— 美元的名义储备水平

P—— 美国价格平减指数

Y—— 名义GNP(美元)

σ_{BP}—— 收支平衡的变化

σ_{EX}—— 汇率的变化

(注：括号内的是 t值。回归分析是依据 28个国家从1976到1985年每年的季度数据，总样本容量为1 120。)

(a) 先验地，你认为各个系数的符号如何？你的预期与结果一致吗？ (b) 解释各个偏斜率系数的意义？ (c) 检验各个偏回归系数的统计显著性 (即，零假设：每个实际或总体回归系数为零)。 (d) 如何检验假设：所有的偏斜率系数同时为零？

8.13 根据英国1950~1966年每年的工资变化百分比 (Y)以及每年的失业率 (X)的数据，得到下面的回归结果：

---

1 参见Mohsen Bahami-Oskooee 与Margaret Malixi的《汇率的流动性与 LDCs对国际储备的需求》(Exchange Rate Flexibility and the LDCs Demand for International Reserves)，《数量经济学》 (Journal of Quantitative Economics)，第4卷，第2期，1988年7月，第317-328页。

$$\hat{Y}_t = -1.428\ 2 + 8.724\ 3\left(\frac{1}{X_t}\right)$$

se=(2.065 7) (2.847 8) r^2=0.384 9

F (1,15)=9.39

(a) 解释系数8.744 3的意义。 (b) 检验假设：估计的斜率系数不为零。你将用什么假设？ (c) 如何用F检验来检验上述假设？ (d) 已知 $\bar{Y}=4.8\%, \bar{X}=1.5\%$，求$Y$的变化率？ (e) 求$Y$对$X$的弹性？ (f) 如何检验假设：实际的 r^2为零。

8.14 表8-9给出了德国1971~1980年间消费者价格指数 Y(1980=100)及货币供给，X(亿德国马克)的数据。

(a) 回归

(1) Y对X (2) LnY对ln$X$ (3) LnY对X (4) Y对ln$X$

(b) 解释各回归结果。 (c) 对每一个模型求 Y对X的变化率。 (d) 对每一个模型求 Y对X的弹性，对其中的一些模型，求均值 Y对均值X的弹性。 (e) 根据这些回归结果，你将选择哪个模型？为什么？

表8-9 德国消费者价格指数(Y)(1980=100)与货币供给(X)*

年份	Y	X	年份	Y	X
1971	64.1	110.02	1980	100.0	237.97
1972	67.7	125.02	1981	106.3	240.77
1973	72.4	132.27	1982	111.9	249.25
1974	77.5	137.17	1983	115.6	275.08
1975	82.0	159.51	1984	118.4	283.89
1976	85.6	176.16	1985	121.0	296.05
1977	88.7	190.80	1986	120.7	325.73
1978	91.1	216.20	1987	121.1	354.93
1979	94.9	232.41			

资料来源：International Economic Conditions，Annual Ed，June 1988，The Federal Reserve Bank of St.Louis，p.24.

*单位是德国马克。

8.15 根据下面的数据估计模型：

$$\frac{1}{Y_i} = B_1 + B_2 X_i + u_i$$

Y	86	79	76	69	65	62	52	51	51	48
X	3	7	12	17	25	35	45	55	70	120

(a) 解释B_2的含义。 (b) 求Y对X的变化率？ (c) 求Y对X的弹性？ (d) 用相同的数据，估计下面的回归模型

$$Y_i = B_1 + B_2\left(\frac{1}{X_i}\right) + u_i$$

(e) 你能比较这两个模型的 r^2值吗？为什么？ (f) 如何判定哪一个模型更好一些？

8.16 当因变量不同时，比较两个 r^2值。[1] 假定你想要比较 (8-19)增长模型与 (8-24)线性趋势模型的 r^2值。具体过程如下：

(a) 求ln $\hat{Y}_t$，即从模型 (8-19)中求出每个观察值的对数。 (b) 求步骤(a)得到的值的反

1 详细细节和具体的计算可参阅， Damodar N.Gujarati《经济计量学基础》(Basic Econometrics)，第3版，McGraw-Hill,纽约，1995，第209-210页。

对数。 (c) 计算根据步骤 (b)得到的 r^2值与根据习题 (6-5)定义的 r^2值。 (d) 这个 r^2值就可以与线性模型 (8-24)的 r^2值相比较。

8.17 根据表 8-5提供的 GNP/货币供给数据，得到下面的回归结果 (Y=GNP，X=货币供给)：

模 型	截 距	斜 率	r^2
双对数模型	0.551 3	0.988 2	0.992 6
	t=(3.165 2)	(41.889)	
对数-线性模型	6.861 6	0.000 57	0.949 3
(增长模型)	t=(100.05)	(15.597)	
线性-对数模型	-163 29.0	258 4.8	0.983 2
	t=(-23.494)	(27.549)	
线性模型	101.20	1.532 3	0.991 5
(LIV模型)	t=(1.369)	(38.867)	

(a) 解释每一个模型斜率的意义。 (b) 对每一个模型，估计 GNP对货币供给的弹性，并解释该弹性。 (c) 所有的 r^2值可直接比较吗？如果不能，哪些可以直接比较？ (d) 你将选择哪个模型？在选择模型时，你考虑的标准有哪些？ (e) 货币学家认为，货币供给的变化率与 GNP之间存在一一对应的关系。上述结果证实了这个观点吗？你将如何验证？

8.18 参考表 8-3给出的能源需求数据。我们用线性模型拟合数据，而不是对数 -线性模型：

$$Y_t=B_1+B_2X_{2t}+B_3X_{3t}+u_t$$

(a) 估计回归系数，求标准差， R^2以及校正的 R^2。 (b) 解释各回归系数。 (c) 估计的各个偏回归系数是统计显著的吗？用 p值回答这个问题。 (d) 建立 ANOVA表，并检验假设：$B_2=B_3=0$。 (e) 计算收入弹性和价格弹性 (用 Y，X_2，X_3的均值)。这些弹性如何才能与式(8-12)给出的收入和价格弹性相比较？ (f) 按照习题 8.16的步骤，比较线性模型和对数-线性模型的 R^2值。 (g) 对上面的变量线性模型的残差作正态概率图。你得出什么结论？ (h) 对对数-线性模型的残差作正态概率图，判断残差是否近似正态分布。 (i) 如果(g)和(h)的结论不同，你将选择哪个回归模型，为什么？

8.19 美国互助基金对投资顾问 (见表 8-10)：

(a) 你将选择哪种回归模型？为什么？ (b) 你所选择的模型有哪些特征？

表8-10 互助基金的管理费用计划

费用 Y(%)	资产净值 X/(亿美元)	费用 Y(%)	资产净值 X/(亿美元)
0.520 0	0.5	0.411 5	30.0
0.508 0	5.0	0.402 0	35.0
0.486 4	10.0	0.394 4	40.0
0.460 0	15.0	0.388 0	45.0
0.439 8	20.0	0.382 5	55.0
0.423 8	25.0	0.373 8	60.0

8.20 为了解释在大的商业银行的商业贷款行为， Bruce J.Summers 运用下面的模型 [1]：

$$Y_t=\frac{1}{A+Bt} \tag{1}$$

其中，Y表示商业或工业贷款 (C&I)，单位为百万美元，按每月计量。分析的数据

1 参见 Bruce J.Summers，“ A Time Series Analysis of Business Loans at Large Commercial Banks，” *Economic Review*，Federal Reserve Bank of St.Louis，May/June，1975，pp.8-14.

从1966年到1967年的月数据，共有 24个观察值。

作者为了作回归分析，用了下面模型：

$$\frac{1}{Y_t} = A + Bt \tag{2}$$

$$\frac{1}{Y_t} = 52.00 - 0.20t \tag{3}$$

$$t=(96.13)(-24.52) \qquad \bar{R}^2 = 0.84$$

$$\frac{\widehat{1}}{Y_t} = 26.179 - 0.14t \qquad \text{DW=0.04*} \tag{4}$$

$$t = (196.70)\ (-66.52) \qquad \bar{R}^2 = 0.97 \qquad \text{DW=0.03*}$$

*表示杜宾-瓦特森(D-W)统计量(参见第12章)。

(a)为什么用模型(2)而不用模型(1)？ (b)这两个模型有什么性质？ (c)解释模型(3)和模型(4)的斜率的含义。它们是统计显著的吗？ (d)怎样求这两个回归方程截距和斜率的标准差？ (e)在C&I活动中，纽约银行和非纽约银行的行为有所不同吗？如何检验这种差别，如果可以的话，写出正规的检验步骤。

附录8A 对 数

考虑数字5和25。我们知道

$$25=5^2 \tag{8A-1}$$

我们称指数2是以5为底25的对数。

若

$$Y=b^x(b>0) \tag{8A-2}$$

则，

$$\log_b Y=X \tag{8A-3}$$

式(8A-2)成为指数函数，式(8A-3)成为对数函数。因此，指数函数和对数函数可以相互转换。

虽然任何一个正数都可以作底数，但在实际中，用的最为广泛的是以 10和e=2.718 28...为底数。

以10为底的对数成为常用对数。

$$\log_{10}100=2, \qquad \log_{10}30\approx1.48$$

也即，在前一种情况下， $100=10^2$，在后一种情况下， $30=10^{1.48}$。

以e为底的对数成为自然对数。

$$\log_e100\approx4.601\ 5, \qquad \log_e30\approx3.401\ 2$$

所有这些都可用计算器来计算。

按照惯例，以10为底的对数常用符号 log表示，以 e为底的对数常用符号 ln表示。因此，上面各例，我们可以写为 log100, log30或ln100, ln30。

常用对数和自然对数之间有一个固定的关系：

$$\ln X=2.3026\log X \tag{8A-4}$$

即X的自然对数等于其常用对数乘以 2.302 6。于是，

$$\ln30= 2.302\ 6\log30=2.302\ 6(1.48)=3.401\ 2(\text{近似值})$$

因此，是用常用对数还是自然对数并没有什么关系。但是在数学中，常用以 e为底的

自然对数。因此，在本书中出现的对数均为自然对数，除非有特别的说明。当然，我们可以根据式(8A-4)作相应的转换。

需要牢记的是：没有定义负数的对数。因而，　log(-5),ln(-5)没有意义。

对数的一些性质如下：若 A，B是任意两个正数，可以证明：

1. $\ln(AB)=\ln A+\ln B$　　(8A-5)

即，A，B乘积的对数等于 A的对数与 B的对数之和。

2. $\mathrm{Ln}(A/B)=\ln A-\ln B$　　(8A-6)

即，A与B商的对数等于 A的对数与 B的对数之差。

3. $\ln(A\pm B)\neq \ln A\pm \ln B$　　(8A-7)

即，A与B的和或差的对数不等于 A的对数与 B的对数的和或差。

4. $\ln(A^k)=k\ln A$　　(8A-8)

即，A的K次方的对数等于 A的对数的 K倍。

5. $\ln e=1$　　(8A-9)

即，e的自然对数为 1(就像10的常用对数为 1一样)。

6. $\ln 1=0$　　(8A-10)

即，1的自然对数为 0(1的常用对数也为 0)。

7. 若$Y=\ln X$, 则　$$\frac{dY}{dX}=\frac{1}{X}$$　　(8A-11)

即Y对X的导数等于 $1/X$。

图8A-1描绘了指数函数和对数函数。

$0<Y<1$　那么 $\ln Y<0$

$Y=1$　那么 $\ln Y=0$

$Y>1$　那么 $\ln Y>0$

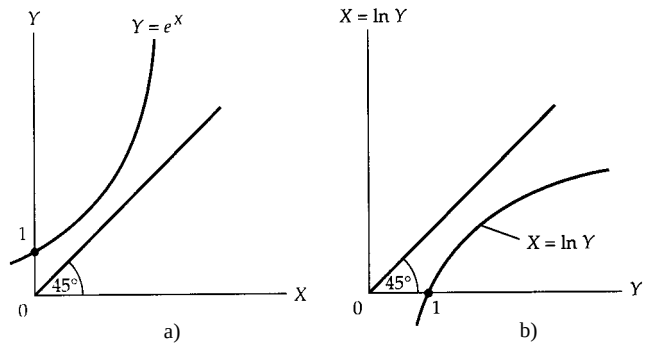


图8A-1　指数和对数函数：
(a)指数函数　(b)对数函数

虽然只有正数才有对数，但是对数值可正可负。很容易证明：

$0<Y<1$，则$\ln Y<0$；$Y=1$，则$\ln Y=0$；$Y>1$，则$\ln Y>0$

虽然图8A-1(b)中的对数曲线正向倾斜的 (表明数值越大，对数值越大)，但是曲线却以一个递减的比率增长 (用数学的语言，其二阶导数为负)。因而，$\ln(10)=2.302\ 6$(近似值),$\ln(20)=2.995\ 7$(近似值)。也就是说，数值变为两倍，但对数值并不是两倍。

这就是为什么把对数变换称为非线性变换的原因。这也可以从式 (8A-11)中得到证明，若$Y=\ln X$, $dY/dX=1/X$。也即导数不是一个常量 (回顾变量之间是线性的定义)。

第 9 章

包含虚拟变量的回归模型

到目前为止，在我们所考虑的线性回归模型中，解释变量 X都是定量变量。但是，情况并非总是如此，有些时候，解释变量却是定性的变量；我们把这类定性变量称为虚拟变量。这些变量还有其他的一些名称，比如，指标变量、二元变量、分类变量、二分变量。本章我们将介绍如何将虚拟变量引入模型并使模型更加丰富和完善。

9.1 虚拟变量的性质

通常在回归分析中，应变量不仅受一些定量变量的影响 (比如，收入、产出、成本、价格、重量、温度等)，而且还受一些定性变量的影响(比如，性别、种族、肤色、宗教、民族、罢工、政团关系、婚姻状况)。举个例子，一些研究者在研究报告中指出，在其他条件不变的情况下，女大学教师的收入比相应的男教师的收入低，类似地，女学生的 S.A.T.的数学平均分数比相应的男生低。(参见表5-5。)无论这种差别的原因如何，在遇到这类问题时，都应该把性别这个定性变量作为解释变量包括到模型之中。当然，我们还可列举其他的一些例子。

这样的定性变量通常表明了具备或不具备某种性质，比如，男性或女性，黑人或白人，天主教徒或非天主教徒，公民或非公民。把这些定性因素“ 定量化 ”的一个方法是建立人工变量，并赋值 0和1，0表示变量不具备某种属性， 1表示变量具备某种属性。例如， 1可代表男性，0代表女性；1代表某人是大学毕业， 0代表某人不是大学毕业； 1代表某人是民主党成员，0代表某人是共和党成员，如此等等；我们将这类取值为 0，1的变量称为 虚拟变量 (dummy variable)。我们用符号 D表示虚拟变量，而不是常用的符号 X。用以强调该变量是定性的变量。

虚拟变量与定量变量一样可用于回归分析。事实上，一个回归模型的解释变量可以仅仅是虚拟变量。解释变量仅是虚拟变量的模型称为 方差分析模型 (analysis-of-variance models)(ANOVA)。我们来看下面的一个例子：

$$Y_i=B_1+B_2D_i+u_i \tag{9-1}$$

其中 Y=初职年薪

$$D_i=\begin{cases}1 & \text{，大学毕业}\\ 0 & \text{，其他(比如，非大学毕业)}\end{cases}$$

模型(9-1)与我们前面讨论过的双变量模型类似，只是这里的解释变量不是定量变量 X，而是虚拟变量D；上面已经讲过，从现在起，用 D表示虚拟变量。

假定随机扰动项满足古典线性回归模型(CLRM)的基本假定，根据模型(9-1)得到：[1]

非大学毕业生的初职年薪的期望为：

$$E(Y_i \mid D_i=0)=B_1+B_2(0)=B_1 \tag{9-2}$$

大学毕业生的初职年薪的期望为：

$$E(Y_i \mid D_i=1)=B_1+B_2(1)=B_1+B_2 \tag{9-3}$$

从这些回归中可以看出，截距 B_1表示非大学毕业生的平均初职年薪，“斜率”系数 B_2表明大学毕业生的平均初职年薪与非大学生的差距是多少； (B_1+B_2)表示大学毕业生的平均初职年薪。

用普通最小二乘法很容易检验零假设：大学教育没有任何益处 (即$B_2=0$)，并可根据t检验值判定b_2是否是统计显著的。

例9.1 大学毕业生和非大学毕业生的初职年薪

表9-1给出了10个不同教育水平的人的初职年薪数据。

表9-1 初职年薪数据

初职年薪Y /千美元	教育(1=大学教育, 0=非大学教育)	初职年薪Y /千美元	教育(1=大学教育, 0=非大学教育)
21.2	1	18.5	0
17.5	0	21.7	1
17.0	0	18.0	0
20.5	1	19.0	0
21.0	1	22.0	1

模型(9-1)OLS回归结果如下：

$$\begin{aligned} \hat{Y}_i &= 18.00+3.28D_i \\ \text{se} &= (0.31) \quad (0.44) \\ t &= (57.74) \quad (7.444) \quad r^2=0.873\ 7 \\ p\text{值} &= (0.000) \quad (0.000) \end{aligned} \tag{9-4}$$

从回归结果可以看出，估计的非大学毕业生的平均初职年薪为 18 000美元$(=b_1)$，大学毕业生的平均初职年薪为21 280美元(b_1+b_2)。根据表9-1提供的数据，很容易计算出非大学毕业生和大学毕业生的平均年薪分别为 18 000美元和21 280美元，正好与它们的估计值相同。

根据括号中的t值，很容易验证b_2是统计显著的(也即显著不为零)，表明非大学毕业生和大学毕业生的初职年薪有差距。实际上也确实如此，学历越高，收入越高！

图9-1描绘了式(9-4)的回归结果。从图中可以看出，回归函数是一个分段函数——非大学毕业生的平均初职年薪是 18 000美元，而大学毕业生的平均初职年薪突升到21 280美元(上升了3 280美元)；两类不同受教育水平的个人年薪在其各自的均值附近波动。

1 由于虚拟变量通常取0和1，所以它是非随机的；也就是说，虚拟变量的取值是固定的。而且由于我们一直假设解释变量X是非随机的，因此就模型（9-1）而言，一个或若干个解释变量是虚拟变量并不会带来任何特殊的问题。简言之，虚拟变量不会有新的估计问题，我们可以用普通最小二乘法对包含虚拟变量的模型的参数进行估计。

例9.2 工作权利法对工会会员的影响

为了研究工作权利法的效果(该法使工会的劳资谈判合同合法化),Brennan等人[1]建立了工会会员(属于工会的工人占所有工人的百分比)对工作权利法(1980年)的函数模型。这项研究包括了50个州,其中19个州制定了工作权利法,31个州允许有工会会员制度(即工会允许进行劳资谈判)。回归结果如下:

$$\widehat{Y}_i = 26.68 - 10.51D_i$$

$$\text{se}=(1.00)\qquad (1.58)\qquad r_2=0.497\ 0 \qquad (9\text{-}5)$$

$$t=(26.68)\qquad (6.65)$$

$$p\text{值}=(0.000)\qquad (0.000)$$

其中,

Y —— 工会成员占工人的比例(1980)

$D=\begin{cases}1, & \text{制定工人工作权利法的州}\\ 0, & \text{未制定工人工作权利法的州}\end{cases}$

从回归结果可以看出,在实施工人工作权利法的州中,工会会员占工人总数的比例为26.68%,而在未实施工人权利法的州中,工会会员占工人总数的比例为16.17%(26.68 - 10.51)。这个差别是显著的,因为虚拟变量的系数显著不为零(验证这一点)。这些结果表明在实施工人权利法的州的工会会员占工人总数的比例比未实施工人权利法的州的比例高,这也并不值得奇怪。

在社会学、心理学、教育学及市场研究等领域,形如(9-4)、(9-5)的ANOVA模型用得很广泛,而在经济学中一般用得很少。在许多的经济研究中,回归模型中的解释变量有些是定量的,有些是定性的。我们将这种回归模型称为协方差模型(ANCOVA)。在本章后面的部分,我们将主要讨论这类模型。

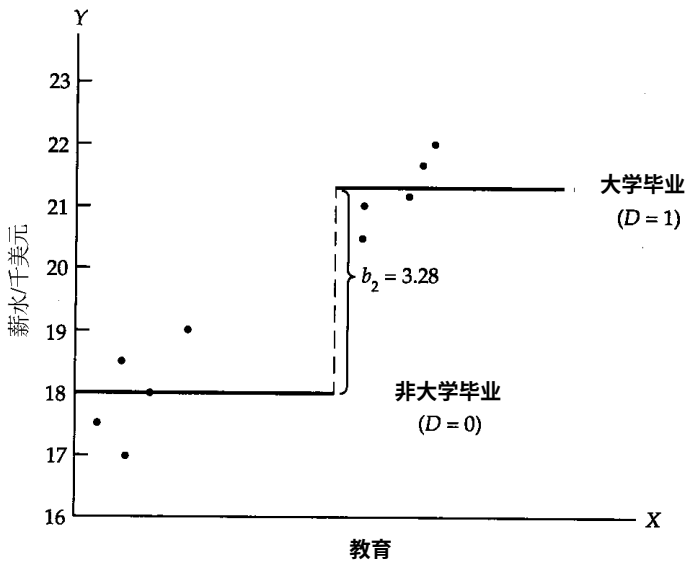


图9-1 不同教育水平下的初职年薪

1 参见Michael J Brennan and Tomas M.Carroll,*Preface to Quantitative Economics and Econometrics*, 4th ed, South-Western Publishing, Cincinnati, Ohio, 1987.符号稍有变动。

9.2　包含一个定量变量，一个两分定性变量的回归模型

考虑下面的一个ANCOVA模型：

$$Y_i=B_1+B_2D_i+B_3X_i+u_i \tag{9-6}$$

其中　Y_i —— 大学教师的年薪　　X_i —— 教龄

$D_i=\begin{cases}1 & \text{，男教师}\\ 0 & \text{，女教师}\end{cases}$

模型(9-6)包含了一个定量的变量 X(教龄)和一个定性变量(性别)(有两类，男教师和女教师)。

对模型(9-6)解释如下。与平常一样，假定 $E(u_i)=0$，则，

男教师平均年薪：

$$E(Y_i \mid X_i，D_i=0)=B_1+B_3X_i) \tag{9-7}$$

女教师平均年薪：

$$E(Y_i \mid X_i，D_i=1)=(B_1+B_2)+B_3X_i) \tag{9-8}$$

图9-2描绘了这两种不同的情况。(为了说明的方便，假定 $B_1>0$)

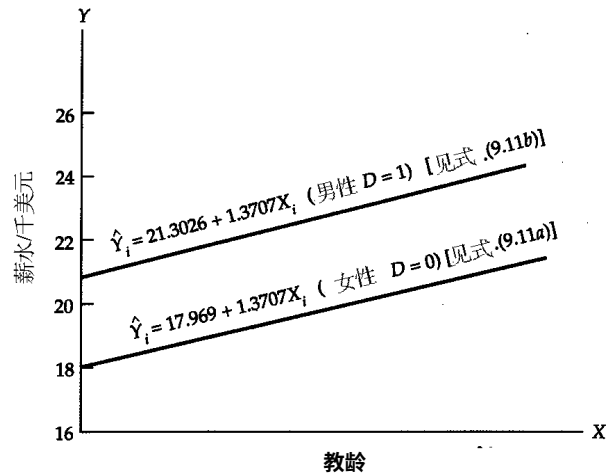


图9-2　根据表9-2数据的回归结果

总之，模型(9-6)表明男、女教师的平均年薪对教龄的函数具有相同的斜率 (B_3)，但截距不同。换句话说，男教师的平均年薪水平与女教师的平均年薪水平不同 (多了 B_2)，但男、女教师平均年薪对教龄的变化率相同。如果斜率相同的假定是有效的，[1] 那么，很容易检验零假设：回归方程(9-7)和(9-8)有相同的截距(也即没有性别歧视)，并根据 t 检验结果判定 b_2 的统计显著性。如果 t 检验表明 b_2 是统计显著的，我们就可以拒绝零假设：男、女教师的平均年薪水平相同。

在继续这个模型之前，先给出虚拟变量的一些性质：

(1) 为了区别男、女两类的不同，我们仅引入了一个虚拟变量，D_i。$D_i=1$ 代表男性，$D_i=0$ 则表示女性，因为这里只有两种可能的结果。因此，一个虚拟变量足可以区分两个不同的种类。若模型包含截距项，则模型(9-6)可写为：

$$Y_i=B_1+B_2X_i+B_3D_{1i}+B_4D_{1i}+u_i \tag{9-9}$$

1　这个假设的有效性可按照9.6节的步骤进行检验。

其中，Y与X与前面的定义相同，

$$D_{1i}=\begin{cases}1 & \text{，男教师}\\ 0 & \text{，女教师}\end{cases}\qquad D_{2i}=\begin{cases}1 & \text{，女教师}\\ 0 & \text{，男教师}\end{cases}$$

我们无法估计模型(9-9)，因为D_{1i}与D_{2i}存在完全共线性(即完全的线性关系)。为了更清楚地了解这一点，假定有这样的一个样本，该样本包括三个男教师，两个女教师。其数据矩阵如下：

		D_1	D_2	X
男	Y_1	1	0	X_1
男	Y_2	1	0	X_2
女	Y_3 =	1	1	X_3
男	Y_4	1	0	X_4
女	Y_5	1	1	X_5

数据矩阵右边的第一列代表了共同的截距，B_1。现在很容易验证：$D_1=(1-D_2)$或$D_2=1-D_1$，也即D_1，D_2完全共线性。在第7章已经讨论过，当存在完全共线性或多重共线性时，不可能得到参数的惟一估计值。

解决完全共线性问题有许多不同的方法。但是最简单的一种就是按照模型(9-6)那样，指定一个虚拟变量，也即若有定性变量有两不同类，例如，性别，就可以仅将模型引入一个虚拟变量。在这种情况下，数据矩阵中就不会有D_2列，这样就可以避免完全共线性问题。一般的规则是：如果一个定性的变量有m类，则要引进$(m-1)$个虚拟变量。在这个例子中，性别有两类，因而在模型(9-6)中仅引进一个虚拟变量。如果不遵循这个规则，就会陷入虚拟变量陷阱(dummy variable trap)，也即完全多重共线性(perfect multicollinearity)情形。

(2) 虚拟变量的赋值是任意的。在这个例子中，令$D=1$，代表女教师，$D=0$，代表男教师；当然，赋值可根据习惯而定。

(3) 赋值为0的一类常称为基准类(base)，对比类(bench mark)，控制类(control)，对比(comparison)或者遗漏类(omitted category)等等。因此，在模型(9-6)中女教师就是基准类。注意，(共同的)截距B_1是基准类的截距，因为若取$D=0$(仅有男教师)，对(9-6)回归，得到的截距为B_1。同样地，对于基准类的选择也是根据研究的目的而定的。

(4) 虚拟变量D的系数称为差别截距系数(differential intercept coefficient)，因为它表明了取值为1的类的截距值与基准类截距值的差距。因此，在模型(9-8)中，B_2表明了在男教师收入回归方程中的截距与女教师收入回归方程中截距的差别是多少。

例9.3 实例一则：教师年薪与教龄、性别的关系。

为了说明ANCOVA模型，我们来看表9-2中数据。

根据该数据，得到的OLS回归结果如下：

$$\begin{aligned}&=17.969+1.370\ 7X_i+3.333\ 6D_i\\ &\text{se}=(0.191\ 9)\quad(0.035\ 6)\quad(0.155\ 4)\qquad\qquad(9\text{-}10)\\ &t=(93.612\ 0)\quad(38.454)\quad(21.455)\qquad R^2=0.993\ 3\end{aligned}$$

对回归结果的解释如下。当性别变量为常量时，平均年薪将增加 1 371美元。当教龄变量保持不变时，男教师的平均年薪比女教师多 3 334美元。由于虚拟变量的系数是统计显著的(为什么？)，因此我们能够说两类教师的平均年薪不同，虽然男女教师平均年薪对教龄有相同的年增长率， 1 337美元。[1]

1 也有可能两类教师的年薪关于教龄的增长率也不同。在 9.6节中，我们将探讨如何发现实际的情况究竟是怎样的。

(续)

根据(9-10)的回归结果，可以推导出男女教师的平均年薪函数：

女教师平均年薪：

$$\hat{Y}_i = 17.969 + 1.370\ 7X_i \tag{9-11a}$$

男教师平均年薪：

$$\hat{Y}_i = (17.969 + 3.333\ 6) + 1.370\ 7X_i \tag{9-11a}$$
$$= 21.302\ 6 + 1.370\ 7X_i$$

图9-2描绘了上述回归结果。

表9-2　大学教师的年薪、教龄及性别

初职年薪 Y/千美元	教龄 X_2	性别(1=男性, 0=女性)	初职年薪 Y/千美元	教龄 X_2	性别(1=男性, 0=女性)
23.0	1	1	25.0	5	0
19.5	1	0	28.0	5	1
24.0	2	1	29.5	6	1
21.0	2	0	26.0	6	0
25.0	3	1	27.5	7	0
22.0	3	0	31.5	7	1
26.5	4	1	29.0	8	0
23.1	4	0			

注意，两条回归直线是平行的，因为我们 (隐含地)假定了两类教师的平均年薪对教龄的变化率是相同的。所不同的是初职年薪，女教师的为 17969美元，男教师为21303美元，而这两个数值正是上述两条回归直线的两个截距 (截距度量了解释变量取零值时的应变量的平均值)。[1]

例9.4　不同规模报酬对产出的影响

为了研究对于同一资本条件下，企业经营不同的产品是否会带来不同的收益率，J.A.Dalton [2] 根据48个企业的数据，得到如下的回归结果：

$$\hat{Y}_i = 1.339 + 1.490D_i + 0.246X_{2i} - 9.507X_{3i} - 0.016X_{4i}$$
$$\text{se} = (1.380)\quad (0.056)\quad (4.244)\quad (0.017)\qquad R^2 = 0.26 \tag{9-12}$$
$$t = (1.079)\quad (4.285)\quad (-2.240)\quad (-0.941)$$
$$p\text{值} = (0.143\ 3)\quad (0.000)\quad (0.015\ 1)\quad (0.175\ 9)$$

式中　Y_i —— 规模报酬率　　D_i —— 1，企业的产品有差别

X_2 —— 市场份额　　X_3 —— 企业数目　　X_4 —— 行业增长率

回归结果表明，当变量 X保持不变时，存在产品差别的企业的规模报酬率比没有产品差别企业的高 1.49个百分点。但是，由于虚拟变量的系数不是统计显著的 (为什么？)，所以用统计学的语言，我们不能说在 X为常量时，两组企业的平均报酬率水平有任何差别。

1　一般的，截距没有实际的经济意义，但是在某些情况下，截距有特殊的含义；在这个例子中，截距代表了教师的平均初职年薪。

2　参见J.A.Dalton and S.L.Levin, "Market Power:Concentration and Market Share," *Industrial Organization Review*,Vol.5,1977,pp.27-36.符号略有变动。

9.3 虚拟变量有多种分类的情况

假定根据横截面数据，我们想要做个人假期旅游的年支出对其收入与受教育水平的回归。由于教育这个变量是一个定性的变量，因此，假定教育水平有如下几等：未达到中学水平，中学水平，大学水平。与前面的情况不同，这里教育变量有三种分类。因此，根据虚拟变量的个数应比变量的分类数少一个的规则，我们引入两个虚拟变量来表示三种不同的教育水平。

假定教育水平不同的三个群体有相同的斜率，但截距不同，我们用下面的模型：

$$Y_i=B_1+B_2D_{2i}+B_3D_{3i}+B_4X_i+u \tag{9-13}$$

式中 Y_i —— 用于假期旅游的年支出

X_i —— 年收入

$D_2=\begin{cases}1, & \text{中学教育}\\ 0, & \text{其他}\end{cases}$ $D_3=\begin{cases}1, & \text{大学教育}\\ 0, & \text{其他}\end{cases}$

注意，在对上面虚拟变量的赋值中，我们将“未达到中学水平”视为基准类。[1] 因此，截距B_1代表了这一类的截距。差别截距 B_2，B_3表明了其他两类的截距与基准类的截距的差距有多大。假定$E(u)=0$(标准假定)，从(9-13)的回归结果可得：

未达到中学水平的平均旅游支出：

$$E(Y_i \mid D_2=0，D_3=0，X_i)=B_1+B_4X_i \tag{9-14}$$

中学水平的平均旅游支出：

$$E(Y_i \mid D_2=1，D_3=0，X_i)=(B_1+B_2)+B_4X_i \tag{9-15}$$

大学毕业的平均旅游支出：

$$E(Y_i \mid D_2=0，D_3=1，X_i)=(B_1+B_3)+B_4X_i \tag{9-15}$$

图9-3描绘了上述三条回归直线(根据例9.5中的数据)。

对模型(9-13)估计之后，根据t检验的结果，很容易验证差别截距 B_2，B_3各自均是统计显著的(也即显著不为零)。

这里有一个问题：为什么不分别对式(9-14)，(9-15)和(9-16)进行回归，而是对式(9-13)进行回归呢？对式(9-14)，(9-15)和(9-16)进行回归并不存在什么问题。但是，如果对它们进行回归，那么，如何发现这三个回归有什么差别呢？我们随后将讨论这个问题 (参见9.6节)。但对式(9-13)回归的一个优点是能够从这一个回归方程中推导出式 (9-14)及(9-15)和(9-16)。

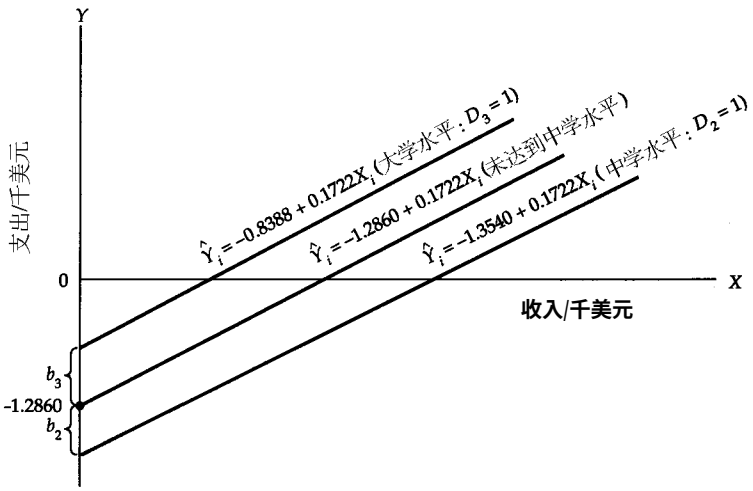


图9-3 旅游支出与不同教育水平下收入的关系

1 当然，其他任何一类都可作为基准类。虽然特殊的研究目标或许表明了哪一类应作为基准类，但是这种选择纯粹是个人的行为。

例9.5　假设一例(旅游支出与收入和教育的关系)

为了解释模型(9-13)，我们来看表9-3给出的数据。根据这些假设的数据得到下面的回归结果：

$$\hat{Y}_i = -1.286\ 0 + 0.172\ 2X_i - 0.068\ 0D_{2i} + 0.447\ 2D_{3i}$$
$$\text{se}=(0.269\ 4)\quad(0.014\ 7)\quad(0.170\ 8)\quad(0.395\ 6) \tag{9-17}$$
$$t=(-4.773\ 8)\quad(11.728\ 0)\quad(-0.398\ 2)\quad(1.130\ 4)\quad R^2=0.996\ 5$$
$$p\text{值}=(0.000)\quad(0.000)\quad(0.349\ 0)\quad(0.141\ 2)$$

表9-3　旅游支出(Y)、收入(X)、受教育水平

Y/千美元	X/千美元	中学教育D_2=1	大学教育D_3=1
6.0	40	0	1
3.9	31	1	0
1.8	18	0	0
1.9	19	0	0
7.2	47	0	1
3.3	27	1	0
3.1	26	1	0
1.7	17	0	0
6.4	43	0	1
7.9	49	0	1
1.5	15	0	0
3.1	25	1	0
3.6	29	1	0
2.0	20	0	0
6.2	41	0	1

注：当$D_2=D_3=0$，观察值表示了未中学毕业。(为什么？)

回归直线见图9-3。

回归结果表明，在其他条件不变时，随着收入的增加，比如说收入增加一美元，平均的旅游支出将增加17美分。我们更关注的是受教育水平对旅游支出的影响。由于在5%的显著水平下，两个虚拟变量均是统计不显著的(检查这个结果)，因而，在收入不变时，受教育水平看似对平均旅游支出没有显著影响。换句话说，三组的截距并没有显著不同。因此，在图9-3中，虽说各个截距的数值不同，但是从统计上说，它们是相同的。

9.4　包含一个定量变量，两个定性变量的回归模型

虚拟变量的技术可以推广到解释变量中有不止一个定性变量的情形。让我们回到大学教师年薪(9.6)一例中，但是现在假定除了教龄、性别以外，肤色也是一个重要的决定因素。为了简便，假定肤色有两种，白种和非白种。我们可将模型(9-6)重写为：

$$Y_i = B_1 + B_2D_{2i} + B_3D_{3i} + B_4X_i + u_i \tag{9-18}$$

式中　Y_i——年薪

X_i——教龄

$D_{2i}=\begin{cases}1, & \text{男教师}\\ 0, & \text{男教师}\end{cases}$　　$D_{3i}=\begin{cases}1, & \text{白种}\\ 0, & \text{非白种}\end{cases}$

注意：性别和肤色这两个定性变量各自均有两类，因此它们每一个都需要一个虚拟变量(以避免虚拟变量陷阱)，这里的基准类是“非白种女教师”。

假定$E(u_i)=0$，则根据模型(9-18)得到不同的平均年薪函数如下：

非白种女教师平均年薪：

$$E(Y_i \mid D_2=0，D_3=0，X_i)=B_1+B_4X_i \tag{9-19}$$

非白种男教师平均年薪：

$$E(Y_i \mid D_2=1，D_3=0，X_i)=(B_1+B_2)+B_4X_i \tag{9-20}$$

白种女教师平均年薪：

$$E(Y_i \mid D_2=0，D_3=1，X_i)=(B_1+B_3)+B_4X_i \tag{9-21}$$

白种男教师平均年薪：

$$E(Y_i \mid D_2=1，D_3=1，X_i)=(B_1+B_2+B_3)+B_4X_i \tag{9-22}$$

再一次地，假定上述回归的截距是不同的，但斜率都相同，为 B_4。

利用OLS法对模型(9-18)进行估计，我们可以根据回归结果检验各种假设。例如，若差别斜率b_3(B_3的估计量)是统计显著的，则意味着肤色对教师薪水有影响。类似地，若差别斜率 b_2是统计显著的，则表明性别也对教师薪水有影响。最终地，若两个差别斜率都是统计显著的，则表示性别和肤色都是教师薪水的重要决定因素。因此，不用对每个方程 (性别和肤色)进行回归就能判定存在那些可能的影响因素，因为可以很容易从 (9-18)推导出这些单个的回归结果。

9.5 模型的推广

读者可能想到，我们可将模型扩展到包括多个定量变量和多个定性变量的情形。但是必须注意的是：每个虚拟变量(相对定性变量)的个数要比该变量的分类数少一。我们来看下面的一个例子。

例9.6 政党对竞选活动的资助[1]

Wilhite 和John Theilmann在研究1982年政党对国会选举的资助中，得到如下回归结果，见表9-4。在这个回归方程中，应变量是 PARTY$(政党对当地侯选人的资助)，$GAP,VGAP和PU是三个定量变量，OPEN，DEMOCRAT和COMM是三个定性变量，每一个定性变量分为两类。

该回归结果说明了什么呢？$GAP越大(也即竞争对手有巨额资助)，政党对当地候选人的资助就越少。VGAP越大，也即竞争对手在以前的竞选中获胜的次数越多，则国会对该候选人的资助就越少(这个预期并不是建立在1982年的选举结果之上的)。公开的竞争可能从国会中吸引更多的资助以确保在国会中的席位，该预期与回归结果一致。政党越忠诚(PU)，得到国会的资助就越多。由于民主党的竞选财力比共和党的少，因此，预期虚拟变量DEMOCTAT的符号为负，事实也的确如此(共和党的竞选资助回归方程的截距比其竞争对手的小)。虚拟变量COMM的符号预期为正，因为如果你也投票选取，而且有幸成国会中分发竞选资金的一员，那么你很可能将资金更多地拨给你所投票的那个政党。

1 选自Wilhite and John Theilmann“Campaign Contributions by Political Parties:Ideology Versus Winning,” *Atlantic Economic Journal* Vol,XVII, June 1989, pp.11-20.

(续)

表9-4 政党资助

解释变量	系 数	解释变量	系 数
$GAP	−8.189*	DEMOCRAT	-9.986*
	(1.863)		(0.557)
VGAP	0.032 1	COMM	1.734*
	(0.022 3)		
OPEN	3.582*	R_2	0.70
	(0.729 3)		
PU	18.189*	F	188.4
	(0.849)		

资料来源：Wilhite and John Theilmann "Campaign Contributions by Political Parties:Ideology Versus Winning." *Atlantic Economic Journal*, Vol.XVII, June1989, pp11-20.p.15.Table2(adapted).

注：括号内的数字是标准差

*表示在0.01的显著水平下是显著的。

$GAP——度量候选人财政

VGAP——在以前竞选中不同政党获得投票的差距

$$\text{OPEN} = \begin{cases} 1, & \text{公开席位竞选} \\ 0, & \text{其他} \end{cases}$$

PU——根据《国会季刊》所计算的政党团结指数

$$\text{DEMOCRAT} = \begin{cases} 1, & \text{民主党成员} \\ 0, & \text{其他} \end{cases} \qquad \text{COMM} = \begin{cases} 1, & \text{民主党成员或共和党成员} \\ 0, & \text{其他(也即非两党成员)。} \end{cases}$$

9.6 回归模型中的结构稳定性：虚拟变量法[1]

回顾第7章的7.12节，我们利用Chow检验判定在1970~1995期间，储蓄-收入关系是否在严重萧条之后于1982年经历了一个结构变动。如果没有结构变动，则可用整个样本区间的数据拟合回归方程(7-54)，但是如果存在结构变动，我们则将用式(7-55)和(7-56)分别对两个时期(1970~1981年，1982~1995年)的数据进行拟合。

事实上，在7.12节中，用Chow检验验证了在样本区间内，储蓄-收入关系之间确实存在一个结构变动。但是，Chow检验并未告诉我们不同的两个回归结果究竟是截距值的不同还是斜率的不同，或是两者均不同。再来看式(7-55)和(7-56)，这里为了方便，我们从新将两式记为：

1970~1981年：

$$Y_t = A_1 + A_2 X_t + u_{1t} \tag{9-23}$$

1982~1995年：

$$Y_t = B_1 + B_2 X_t + u_{2t} \tag{9-24}$$

式中 Y——储蓄

X——收入

1 材料选自作者的论文《用虚拟变量检验两线性回归方程系数之间是否相等：一个注解》及《用虚拟变量检验两线性回归方程系数之间是否相等：一个归纳》，发表于《美国统计学家》(*American Statistician*)，第24卷，1970年，第50-52, 18-21页。

u —— 随机误差项

回归方程(9-23)和(9-24)代表了下面4种可能的结果：(参见图9-4)

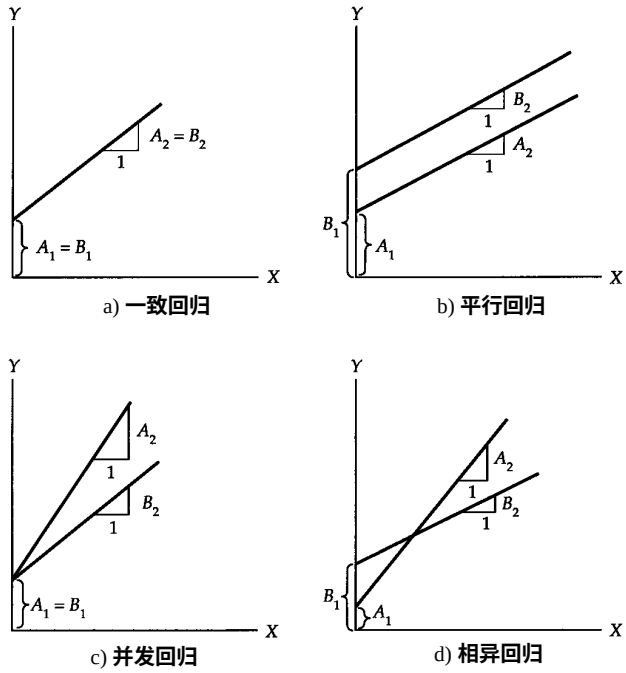


图9-4 两个回归结果可能出现的不同的情况

(1) $A_1=B_1$, $A_2=B_2$; 即两回归相同。我们称之为一致回归。在这种情况下，无须单独建立两个回归方程。

(2) $A_1 \neq B_1$, 但$A_2=B_2$；也即两回归仅仅位置不同(截距不同)。我们称之为平行回归。[参见图9-4(b)]

(3) $A_1=B_1$, $A_2 \neq B_2$；也即两回归截距相同，但斜率不同。我们称之为并发回归。[参见图9-4 (c)]

(4) $A_1 \neq B_1$, $A_2 \neq B_2$；也即两回归完全不同。我们称之为相异回归。[参见图9-4(d)]

在储蓄-收入一例中，如何知道储蓄和收入之间的关系是上述哪种情况呢？虚拟变量技术能够解决这个问题。我们来看下面的回归：

$$Y_t=C_1+C_2D_t+C_3X_t+C_4(D_tX_t)+u_t \tag{9-25}$$

其中 Y —— 个人储蓄

X —— 个人可支配收入

$D_t = \begin{cases} 1, & \text{观察值从1982年开始} \\ 0, & \text{其他(也即观察值直到1981年，见表9-5)} \end{cases}$

为了明确(9-25)回归方程的意义，假定$E(u_t)=0$，得到：

$$E(Y_t \mid D_t=0, X_t)=C_1+C_3X_t \tag{9-26}$$

$$E(Y_t \mid D_t=1, X_t)=(C_1+C_2)+(C_3+C_4)X_t \tag{9-27}$$

式(9-26)和(9-27)分别是萧条前和萧条后的(平均)储蓄函数。换句话说，式(9-26)与式(9-23)相同，式(9-27)与式(9-24)相同：

$$A_1=C_1, \quad A_2=C_3$$

$$B_1=(C_1+C_2), \quad B_2=(C_3+C_4)$$

因此，从单一的回归方程(9-25)可以很容易得到两个不同阶段的回归方程，这再一次证明

了虚拟变量技术的灵活性。

在式(9-25)中，C_2是差别截距(differential intercept)。与前面相同，C_4是差别斜率(differential slope coefficient)，它表明萧条后的储蓄函数与萧条前的储蓄函数的斜率的差距有多大。[1] 注意此时的虚拟变量作乘项，用以区分两个阶段的斜率，这与虚拟变量作加项用以区分两阶段的截距相类似。

例9.7 1970~1995，美国储蓄-收入关系

在表7-6的基础上，我们增加了虚拟变量，见表9-5，根据模型(9-25)，利用表9-5提供的数据得到下面的回归结果：[2]

$$\hat{Y}_t = 1.02 + 152.48D_t + 0.080\ 3X_t - 0.065\ 5(D_tX_t) \quad (9\text{-}28)$$

se=(20.16) (33.08) (0.014 5) (0.015 9)

t=(0.05) (4.61) (5.54) (-4.10)

p值=(0.960) (0.000 0)* (0.000 0)* (0.000 0)*

R^2=0.822

$\bar{R}^2$=0.866

F=54.78

*表示值很小。

表9-5 美国个人储蓄和个人可支配收入

年份	个人储蓄	个人可支配收入	虚拟变量	虚拟变量与DPI的乘积
1970	61.0	727.1	0	0.0
1971	68.6	790.2	0	0.0
1972	63.6	855.3	0	0.0
1973	89.6	965.0	0	0.0
1974	97.6	1 054.2	0	0.0
1975	104.4	1 159.2	0	0.0
1976	96.4	1 273.0	0	0.0
1977	92.5	1 401.4	0	0.0
1978	112.6	1 580.1	0	0.0
1979	130.1	1 769.5	0	0.0
1980	161.8	1 973.3	0	0.0
1981	199.1	2 200.2	0	0.0
1982	205.5	2 347.3	1*	2 734.3
1983	167.0	2 522.4	1	2 522.4
1984	235.7	2 810.0	1	2 810.0
1985	206.2	3 002.0	1	3 002.0
1986	196.5	3 187.6	1	3 187.6
1987	168.4	3 363.1	1	3 363.1
1988	189.1	3 640.8	1	3 640.8
1989	187.8	3 894.5	1	3 894.5
1990	208.7	4 166.8	1	4 166.8
1991	246.4	4 343.7	1	4 343.7
1992	272.6	4 613.7	1	4 613.7
1993	214.4	4 790.2	1	4 790.2
1994	189.4	5 021.7	1	5 021.7
1995	249.3	5 320.8	1	5 320.8

*D=1，1982年开始的观察值。

资料来源：《总统报告》，1997年，数据单位为亿美元，摘自表B-28，第332页。

1 一般地，差别斜率系数告诉我们取值为1的一类的斜率与基准类的斜率的差别有多大。

2 我们假定$\text{var}(u_{1t})=\text{var}(u_{2t})=\text{var}(u_{3t})=\sigma^2$。在Chow检验时，也作同样的假定。

(续)

(9-28)回归结果可用来解释任何其他一个多元回归。读者很容易验证，差别截距和差别斜率各自均是统计显著的 (因为p值很小)，这表明了萧条前和萧条后的储蓄函数明显不同。这一点可从图9-5中清楚地看到(参见图9-4d)。

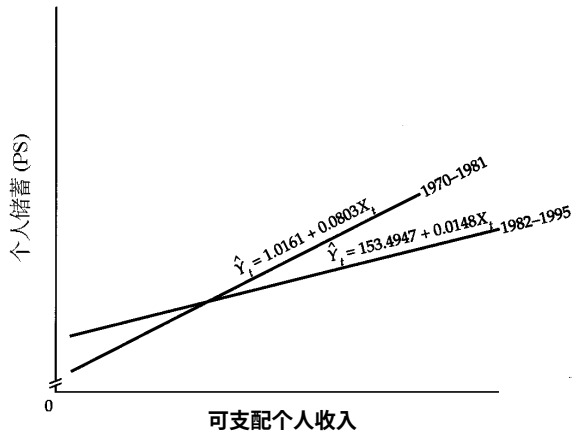


图9-5 美国储蓄-收入关系(1970~1987年)

利用式(9-26)和(9-27)，可推导出两个时期的储蓄函数：

平均储蓄函数：1970~1981年

$$\hat{Y}_t = 1.02 + 0.080\ 3X_t \tag{9-29}$$

平均储蓄函数：1982~1995年

$$\begin{aligned}\hat{Y}_t &= (1.02 + 152.48) + (0.080\ 3 - 0.065\ 5)\ X_t \\ &= 153.5 + 0.014\ 8X_t\end{aligned} \tag{9-30}$$

从回归方程(9-29)和(9-30)中可以看出，1970~1981年期间的截距并不显著不为零；但1982~1995期间的截距显著为正。另一方面，虽然两个时期的斜率(度量了边际储蓄倾向，MPS)都为正而且是统计显著的，但是前一时期的斜率比后一时期的斜率更大：在第一时期，平均而言，每增加一美元收入储蓄将增加8美分；但是在第二时期，平均而言，每增加一美元收入储蓄仅增加了2美分。或许这反映了1982年经济萧条的严重性。

如果将虚拟变量法和Chow检验法做一比较，我们会发现虚拟变量法的一些优点。第一，除了舍入误差外，式(7-55a)和(9-29)，式(7-56a)和式(9-30)几乎相同。第二，与Chow检验法不同，虚拟变量法无须估计三个回归方程(7-54)、(7-55)和(7-56)，只需运行方程(9-28)，因为两个阶段的回归方程可以很容易通过这个单方程推导出来。第三，从(7-55a)和(7-56a)方程中，不易看出这两个回归显著不同；但是，从不同的虚拟变量和截距系数，我们可以指出两回归方程差别的原因。

例9.8 美国菲利普斯曲线如何

为了进一步说明差别截距和差别斜率在检验两回归结果不同时的作用，[1] 回顾式(8-30)描述的菲利普斯曲线(1958~1969年)。后来经济学家发现工资变化率与失业率之间的反向关系不在存在了。一些经济学家甚至认为菲利普斯曲线根本不存在。为了验证事实

1 这个技术可以推广到比较多个回归结果的情形。例如，如果我们要比较三个回归方程，则需引入两个虚拟变量（比较的虚拟变量的分类数少一）。

(续)

是否真的如此，我们将表8-6中的样本数据延续到1977年，则估计的回归方程为：

$$Y_t=B_1+B_2D_t+B_3\left(\frac{1}{X_t}\right)+B_4D_4\left(\frac{1}{X_t}\right)+u_t \tag{9-31}$$

式中 Y —— 小时工资指数的年变化率

X —— 失业率

$D_t=\begin{cases}1, & \text{1969年以前的观察值}\\ 0, & \text{其他(也即从1970~1977年的观察值)}\end{cases}$

回归结果如下：

$$\hat{Y}_t=10.078-10.337D_t-17.549\left(\frac{1}{X_t}\right)+38.137D_t\left(\frac{1}{X_t}\right)$$

$$\begin{aligned}&\text{se}=(1.402\ 4)\quad(1.685\ 9)\quad(8.337\ 3)\quad(9.399\ 9)\\ &t=(7.186\ 0)\quad(-6.131\ 4)\quad(-2.104\ 9)\quad(4.057\ 2)\quad R^2=0.0878\ 7\\ &p\text{值}=(0.000)\quad(0.000)\quad(0.026)\quad(0.000)\end{aligned} \tag{9-32}$$

回归结果表明，差别截距和差别斜率各自均显著不为零，这表明从1969年，菲利普斯曲线有一个显著的可观察到的变化。为了更清楚地看到这一点，我们从式(9-32)推导出两个不同时期的回归结果：

1958~1969年的菲利普斯曲线：

$$\begin{aligned}\hat{Y}_t&=(10.078-10.337)+(-17.549+38.137)\left(\frac{1}{X_t}\right) \qquad (9\text{-}33)=(8\text{-}30)\\ &=-0.259+20.588\left(\frac{1}{X_t}\right)\end{aligned}$$

1970~1977年菲利普斯曲线：

$$\hat{Y}_t=10.078-17.549\left(\frac{1}{X_t}\right) \tag{9-34}$$

特别值得注意的是1970~1977年期间的菲利普斯曲线是正向倾斜的！(注：变量X以倒数形式进入模型，因此在其他条件不变时，斜率应该为正，比如1958~1969年期间的菲利普斯曲线)。这表明失业率越高，小时工资却增加得更快。这一结论无法用经济理论来解释。看起来菲利普斯曲线真的失灵了。[1]

9.7 虚拟变量在季节分析中的应用

许多用月度或季度数据表示的经济时间序列，呈现出季节变化的规律性(季节模式(season patterns))。例如，圣诞前后商店的销售量，节假日期间家庭对货币的需求(现金余额)，夏天对冰激凌、软饮料的需求，假期对旅游的需求等等。通常可以从时间序列中将季节因素或成分除去，这样就可以将注意力集中在其他的成分上，比如说长期趋势成分(稳定的增加或减少)。[2]把季节成分从时间序列从除去的过程称为消除季节成分或季节调整时间序列。美国政府公布的一些重要的经济时间序列数据就是经过季节调整后得到的。

从时间序列中消除季节成分的方法有多种，但我们仅介绍其中的一种方法——虚拟变量法。[3]

1 不仅仅是在美国，现在有大量的经验证据表明在许多工业化国家，菲利普斯曲线都失灵了。

2 一个时间序列包括四个组成部分：季节变动成分，循环变动成分，长期趋势成分，不规则变动成分。

3 其他的方法有：移动平均率法，联系相关法和平均年百分率法等 。有关非技术细节，参见Paul Neebold *Statistics for Business and Economics*,4th ed., Prentice-Hall, Englewood Cliffs, N.J., 1995，Chap.17.

我们用下面的一个例子来说明。

例9.9 澳大利亚支出-消费关系，1977年第一季度~1980第四季度。

表9-6给出了澳大利亚从1977年第一季度到1980第四季度的Y[衣服、硬件、电器、家具的零售价，称为个人消费支出 (PCE)]和X[个人可支配收入 (PDI)]的季度数据。考虑下面的模型：

表9-6 澳大利亚个人消费支出PCE(Y)，PDI(X)*

年份季节	Y	X	D_2	D_3	D_4
1977第一季度	16.63	136.5	0	0	0
第二季度	19.91	132.1	1	0	0
第三季度	19.41	157.5	0	1	0
第四季度	24.01	177.7	0	0	1
1978第一季度	17.55	152.4	0	0	0
第二季度	21.97	150.7	1	0	0
第三季度	20.90	173.0	0	1	0
第四季度	25.61	199.8	0	0	1
1979第一季度	19.46	179.1	0	0	0
第二季度	22.72	167.4	1	0	0
第三季度	22.14	191.6	0	1	0
第四季度	27.42	227.0	0	0	1
1980第一季度	21.42	187.3	0	0	0
第二季度	25.41	185.0	1	0	0
第三季度	25.49	219.2	0	1	0
第四季度	32.07	261.5	0	0	1

*单位为千万美元澳大利亚元

资料来源：New South Wales大学的Eric Sowey博士提供。

$$Y_t=B_1+B_2D_{2t}+B_3D_{3t}+B_4D_{4t}+B_5X_t+u_t \tag{9-35}$$

其中，Y和X的定义与前面的相同，D的定义如下：

$$D_{2t}=\begin{cases}1，\text{第二季度数据}\\0，\text{其他}\end{cases}\quad D_{3t}=\begin{cases}1，\text{第三季度数据}\\0，\text{其他}\end{cases}\quad D_{4t}=\begin{cases}1，\text{第四季度数据}\\0，\text{其他}\end{cases}$$

假定变量“季节”分四类(一年有四个季节)，因此需要引入三个虚拟变量；在这个例子中，令第一季度为基准类：

第一季度的平均消费支出：

$$E(Y_t \mid D_2=0，\ D_3=0，\ D_4=0，\ X_t)=B_1+B_5X_t \tag{9-36}$$

第二季度的平均消费支出：

$$E(Y_t \mid D_2=1，\ D_3=0，\ D_4=0，\ X_t)=(B_1+B_2)+B_5X_t \tag{9-37}$$

第三季度的平均消费支出：

$$E(Y_t \mid D_2=0，\ D_3=1，\ D_4=0，\ X_t)=(B_1+B_3)+B_5X_t \tag{9-38}$$

第四季度的平均消费支出：

$$E(Y_t \mid D_2=0，\ D_3=0，\ D_4=1，\ X_t)=(B_1+B_4)+B_5X_t \tag{9-39}$$

表9-6给出了虚拟变量是如何建立的。

根据模型(9-35)得到回归结果如下：

$$\hat{Y}_t=3.270\ 6+4.212\ 8D_{2t}+1.186\ 6D_{3t}+3.530\ 6D_{4t}+0.094\ 6X_t$$

$$\text{se}=(0.934\ 5)\quad(0.372\ 9)\quad(0.390\ 1)\quad(0.470\ 6)\quad(0.005\ 5)$$

(续)

t=(3.499 8)　(11.296)　(3.041 7)　(7.502 3)　(17.278)　　(9-40)

p值=(0.002)　(0.000)　(0.005)　(0.000)　(0.000)

R^2=0.98 70

回归结果表明，所有的差别截距各自均是统计显著的，这表明不同季节的 Y的平均水平不同，表现出四个明显的“季节”。例如，第四季度的 Y的平均值比第一季度(基准季度)的高3.5(千万)。PDI的系数是统计显著的，而且符号为正，这并没有什么可奇怪的，因为，我们先验地预期PCE与PDI之间是正向关系。

这里需要注意的是，在式(9-40)的回归结果中暗含地假定了季节因素(如果存在的话)仅仅影响截距，而不影响斜率。但是如何知道事实的确如此呢？我们可以用上一节讨论的差别截距和差别斜率法来验证。

我们来看看在本例中上述方法是如何得以应用的。将模型(9-35)扩展为：

$$Y_t=B_1+B_2D_{2t}+B_3D_{3t}+B_4D_{4t}+B_5X_t+B_6(D_{2t}X_t)+B_7(D_{3t}X_t)+B_8(D_{4t}X_t)+u_t \quad (9\text{-}41)$$

仔细观察多重虚拟变量是如何生成的。差别斜率系数 B_6、B_7、B_8，告诉我们第二、第三、第四季度的斜率与基准(也即第一)季度B_5的差别有多大。用模型(9-41)拟合表9-6提供的澳大利亚的数据，得到如下回归结果，见表9-7。

比较回归结果(9-40)和(9-42)，我们发现戏剧性的变化。收入变量PDI的系数在两个回归中都是正的和统计显著的，而且值相等。但是，各差别截距在式(9-40)回归结果中是统计显著的，而在式(9-42)回归结果中却是统计不显著的。不仅如此，在式(9-42)中，没有一个偏斜率是统计显著的。

表9-7　澳大利亚经济模型(9-41)的回归结果

解释变量	系　数	解释变量	系　数	
b_1(常数)	4.452 6	X_t	0.087 4	(9-42)
	t=(1.819 3)*		(5.893 1)*	
D_{2t}	2.408 2	$D_{2t}X_t$	0.011 1	
	(0.693 32)		(0.520 84)	
D_{3t}	-0.423 36	$D_{3t}X_t$	0.009 5	
	(-0.122 13)		(0.806 7)	
D_{4t}	2.289 1	$D_{4t}X_t$	0.007 5	
	(0.708 19)		(0.423 37)	
			R^2=0.987 5	

注：括号内给出的是t值

*表示在5%的显著水平下是统计显著的(单边检验)。

一个重要的问题是：我们相信哪个结果呢——是式(9-40)还是(9-42)？模型(9-40)是直接依据假定：四个季度的截距(而不是斜率)不同。换句话说，这个回归结果是建立在假定四条回归直线(每个季度一条)是平行的基础上的(参见图9-4b)。而另一方面，模型(9-42)更一般地假定所有的回归线是不相似的(见图9-4b)。如果回归结果(9-42)为真，而我们却用了模型(9-40)，那么就犯了模型设定偏差(model specification error 或者 model specification bias)，也即用错误的模型拟合了数据。这种设定偏差的后果可能是严重的。(更多的分析见第13章)在实践中，最好用一般的模型(9-42)。如果所有的差别斜率均是统计不显著的，那么可用模型(9-40)看看差别截距是否是统计显著的。当然，如果各个差别截距是统计不显著的，那么基准类与其他类就没有

差别，也就是一致回归的情况，见图 9-4a。

按照这种方法，综合得出的结论是：四个季度的平均 PCE对PDI的变化率是相同的，但是四个季度的平均PCE的水平是不同的，见图9-6。简言之，我们应选择模型(9-40)。

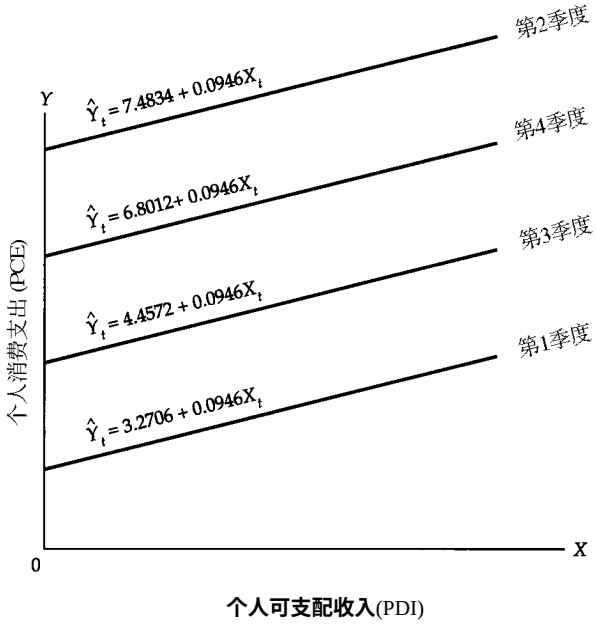


图9-6 澳大利亚PCE-PDI关系图

9.8 小结

本章我们介绍了取值为0、1的虚拟变量是如何引入到回归模型中的。我们通过不同的例子说明了虚拟变量从本质上说是“数据分类器”，它根据样本的属性(性别、婚姻状况、种族、宗教等等)将样本分为各个不同的子群体并对每个子群体进行回归分析。各个子群体的应变量对解释变量(定性变量)的不同反应表现为各子群体的截距或斜率系数存在差别。

虽然虚拟变量技术非常有用，但在使用时仍许谨慎。第一，如果回归模型包含了一个常数项(大多数模型都包含常数项)，那么虚拟变量的个数必须比每个定性变量的分类数少一。第二，虚拟变量的系数必须与基准类相关 —— 取值为零的一类。最后，若模型包含多个定性变量，而且每个定性变量有多种分类，则引入模型的虚拟变量将消耗大量的自由度。因此，应当权衡进入模型的虚拟变量的个数以免超过样本观察值的个数。

本章我们还讨论了模型设定偏差的情况，也即用错误的模型拟合了样本数据。如果各个子群体的截距和斜率都不同，那么我们应当建立一个斜率和截距均不同的模型。在这种情况下，如果仅仅是截距不同将很可能导致设定误差。因此，在具体的分析中经验是至关重要的，尤其在缺乏足够的经济理论指导的情况下。有关设定误差的进一步讨论见第 13章。

习题

9.1 简要解释下列概念

(a) 分类变量 (b) 定性变量 (c) 方差分析模型(ANOVA) (d) 协方差模型(ANOCVA) (e) 虚拟变量陷阱 (f) 不同截距的虚拟变量 (g) 不同斜率的虚拟变量

9.2 下面的变量是定性的还是定量的？

(a) 美国国际收支差额 (b) 政党联盟 (c) 美国对中国的出口 (d) 联合国会员 (e) CPI(消费者价格指数) (f) 教育 (g) 欧共体公民 (h) 关贸组织成员 (i) 美国国会成员 (j) 社会保障受益者

9.3 如果现在有月度数据，在对下面的假设进行检验时，你将引入几个虚拟变量？

(a) 一年中的每月均呈现季节波动趋势。 (b) 只有2月、4月、6月、8月、10月、12月呈现季节波动的趋势。

9.4 在估计下面的模型中，你预见会出现什么问题？

(a) $Y_t=B_0+B_1D_{1t}+B_2D_{2t}+B_3D_{3t}+B_4D_{4t}+u_t$

其中，$D_{it}=\begin{cases}1, & \text{第}i\text{季的观察值，}i=1,2,3,4\\ 0, & \text{其他}\end{cases}$

(b) $\text{GNP}_t=B_1+B_2M_t+B_3M_{t-1}+B_4(M_t-M_{t-1})+u_t$

其中，GNP_t=第t期国民生产总值

M_t=第t期的货币供给

M_{t-1}=第$t-1$期的货币供给

9.5 判断正误并说明理由。

(a) 在模型$Y_i=B_0+B_1X_i+B_2D_i+u$中，令D_i取值(0，2)而不是(0，1)，那么B_2的值将二等分，t值也将二等分。 (b) 引入虚拟变量后，用普通最小二乘法得到的估计量只有当大样本时才是无偏的。

9.6 考虑下面的模型：

$$Y_i=B_0+B_1X_i+B_2D_{2i}+B_3D_{3i}+u_i$$

其中 Y——MPA毕业年收入

X——工龄

$D_2=\begin{cases}1, & \text{哈佛MBA}\\ 0, & \text{其他}\end{cases}$ $D_3=\begin{cases}1, & \text{沃顿MBA}\\ 0, & \text{其他}\end{cases}$

(a) 你预期各系数的符号如何？ (b) 如何解释截距B_2，B_3？ (c) 若$B_2>B_3$，你得出什么结论？

9.7 继续上题，但现考虑下面的模型：

$$Y_i=B_0+B_1X_i+B_2D_{2i}+B_3D_{3i}+B_4(D_{2i}X_i)+B_5(D_{3i}X_i)+u_i$$

(a) 这个模型与习题9.6的模型有什么区别？ (b) 如何解释截距B_4，B_5？ (c) 若B_4，B_5各自均是统计显著的，你将选择这个模型还是习题 9.6的模型？如果不是，你会犯哪类错误？

*(d) 如何检验假设：$B_4=B_5=0$？(选做题，参见第14章的有限的OLS法)

9.8 H.C.Huang,J.J.Siegfried和F.Zardonshty [1] 根据美国1961第一季度~1977第二季度的季度数据估计了对咖啡的需求函数如下：(括号内的数字为t值)

$$\ln\hat{Q}_t=1.278\ 9-0.164\ 7\ln P_t+0.515\ 5\ln I_t+0.148\ 3\ln P'_t$$
$$t=(-2.14)\qquad(1.23)\qquad(0.55)$$
$$-0.008\ 9T-0.096\ 1D_{1t}-0.157\ 0D_{2t}-0.009\ 7D_{3t}\qquad R^2=0.80$$
$$t=(-3.36)\quad(-3.74)\quad(-6.03)\quad(-0.37)$$

1 H.C.Huang,J.J.Siegfried和F.Zardoshty《美国对咖啡的需求， 1963-1977》，《经济学和商业季刊》(*Quarterly Review of Economics and Business*)1980年8月，第36-50页。

式中，Q——(按人口)平均消费咖啡量

P ——每磅咖啡的相对价格(以1967年为不变价)

I——(按人口)平均PDI，单位为美元(以1967年为不变价)

P' ——每磅茶的相对价格(以1967年为不变价)

T——时间趋势，T=1(1961年第一季度)至T=66(1977年第二季度)

D_1——1，第一季度

D_2——1，第二季度

D_3——1，第三季度

ln——自然对数

(a) 如何解释P，I，P'的系数。 (b) 咖啡的需求对价格是富有弹性的吗？ (c) 咖啡和茶是互补品还是替代品？ (d) 如何解释T的系数？ (e) 求美国咖啡消费的增长率或衰减率？如果对咖啡的消费有一个下降的趋势，你将对此作何解释？ (f) 求对咖啡需求的收入弹性？ (g) 如何检验假设：对咖啡需求的收入弹性显著不为1？ (h) 在这个例子中，虚拟变量代表了什么？ (i) 如何解释模型中的虚拟变量？ (j) 哪些虚拟变量是统计显著的？ (k) 美国咖啡消费是否存在明显的季节变动趋势？如果存在的话？如何解释？ (l) 在这个例子中，基准类是什么？如果你选择其他的基准类，结果会有什么变化？ (m) 上述模型仅仅引入了差别斜率虚拟变量，这里隐含的假定是什么？ (n) 假定某人认为这个模型设定有误，因为该模型假定了各个变量的斜率在不同季节保持不变。你将如何修该这个模型从而考虑到差别斜率虚拟变量？ (o) 如果有具体数据的话，你将如何用公式表示出咖啡的需求函数？

9.9 Paul W.Bauer 和 Thomas J.Zlatoper在研究决定开往Cleveland的直接机票的因素中得到下面的回归结果(表形式)用以解释单程头等舱、二等舱和经济舱机票(因变量是单程机票)的价格。解释变量定义如下：

Carriers——载客人数

Pass——总乘客人数

Miles——从出发地到Cleveland的距离

Pop——出发地人口

Inc——出发地人均收入

Corp——潜在商业交通代理

$$\text{Slot}\begin{cases}1, & \text{虚拟变量}\\ 0, & \text{其他}\end{cases}$$

Stop=

$$\text{Meal}\begin{cases}1, & \text{供餐}\\ 0, & \text{其他}\end{cases}\qquad \text{Hub}\begin{cases}1, & \text{若出发地有中转站}\\ 0, & \text{其他}\end{cases}\qquad \text{EA}\begin{cases}1, & \text{向东的航线}\\ 0, & \text{其他}\end{cases}\qquad \text{CO}\begin{cases}1, & \text{洲际航线}\\ 0, & \text{其他}\end{cases}$$

(a) 在这个模型中，引入变量载客人数和载客人数的平方为解释变量的理论依据是什么？载客人数符号为负和载客人数平方符号为正表明了什么？ (b) 引入距离和距离的平方为解释变量的理论依据是什么？观察到的这些变量符号有经济意义吗？ (c) 观察到的人口变量的符号为负，这有什么含义？ (d) 为什么在所有的回归结果中人均收入变量的符号均为负？ (e) 为什么“stop”变量在“头等舱”和“二等舱”回归方程中的符号为正，而在“经济舱”回归方程中符号为负。 (f) 虚拟变量“洲际航线”的符号始终为负。这表明了什么？ (g) 估计每个回归系数的显著性。注：由于观察值的个数足够大，因此在5%的显著水平下，可用正态分布近似t分布。分别利用单边和双边检验。 (h) 为什么虚拟变量“Slot”仅仅在“经济舱”回归方程中是统计显著的？ (i) 由于“头等舱”“经济舱”的观察值的个数相同，均为323个，

能否将它们加总起来 (646个)作一个回归方程？如果可以，如何区别“二等舱”和“经济舱”的观察值(提示：用虚拟变量)。 (j) 对上表中的回归结果作出评价。

解释变量	头等舱	二等舱	经济舱
Carriers	-19.50	-23.00	-17.50
	*t=(-0.878)	(-1.99)	(-3.67)
Carriers^2	2.79	4.00	2.19
	(0.632)	(1.83)	(2.42)
Miles	0.223	0.277	0.079 1
	(5.13)	(12.00)	(8.24)
Miles^2	-0.000 009 7	-0.000 052	-0.000 014
	(-0.495)	(-4.98)	(-3.23)
Pop	-0.005 98	-0.001 14	-0.000 868
	(-1.67)	(-4.98)	(-1.05)
INC	-0.001 95	-0.001 87	-0.004 11
	(-0.686)	(-1.06)	(-6.05)
Corp	3.62	1.22	-1.06
	(3.45)	(2.51)	(-5.22)
Pass	-0.000 818	-0.000 275	0.853
	(-0.771)	(-0.527)	(3.93)
Stop	12.50	7.64	-3.58
	(1.36)	(2.13)	(-2.60)
Slot	7.13	-0.746	17.70
	(0.299)	(-0.067)	(3.82)
Hub	11.30	4.18	-3.50
	(0.90)	(0.81)	(-1.62)
Meal	11.20	0.945	1.80
	(1.07)	(0.177)	(0.813)
EA	-18.30	5.80	-10.60
	(-1.60)	(0.775)	(-3.49)
CO	-66.40	-56.50	-4.17
	(-5.72)	(-7.61)	(-1.35)
常数项	212.00	126.00	113.00
	(5.21)	(5.75)	(12.40)
R^2	0.863	0.871	0.799
观察值个数	163	323	323

资料来源：Paul W.Bauer 和 Thomas J.Zlatoper《经济评论》(Economic Review)，Cleveland联邦储备银行，第25卷，第一期，1989年，表2，3，4，第6-7页。

9.10 在51个学生(其中男生36人，女生15人)的体重(W)对身高(H)的回归分析中，得到下面结果：[1]

1. $\widehat{W}_i = -232.065\ 51 + 5.566\ 2H_i$

   t= (-5.206 6) (8.624 6)

2. $\widehat{W}_i = -122.962\ 1 + 23.823\ 8S_iD + 3.740\ 2H_i$

   t= (-2.588 4) (4.014 9) (5.161 3)

3. $\widehat{W}_i = -107.950\ 8 + 3.510\ 5H_i + 2.007\ 3S_iD + 0.326\ 3H_iD$

   t= (-1.226 6) (2.608 7) (0.018 7) (0.203 5)

其中体重的单位为磅，身高的单位为英寸。

1 Alert Zuker收集了数据并作了各种不同的回归分析。

S(性别) = $\begin{cases}1 & \text{，男生}\\ 0 & \text{，女生}\end{cases}$

D(虚拟变量)——乘积或差别斜率虚拟变量

(a) 你将选择哪个回归，1还是2，为什么？ (b) 如果实际较为理想的回归是2，而你却选择了1，你犯了哪类错误？ (c) 回归2中的性别这个虚拟变量表明了什么？ (d) 模型2中，差别截距是统计显著的，但在模型 3中，差别斜率却是统计不显著的？如何解释这种变化？ (e) 在模型2与模型3之间，你选择哪一个？为什么？ (f) 在模型2与模型3中，变量身高的系数几乎相等，但性别这一虚拟变量的系数相差很大。对此你有什么想法？

为了回答问题(d)，(e)，(f)，这里给出了下面的相关矩阵。

解释：例如，身高和性别的相关系数是0.627 6，性别和交互虚拟变量的相关系数是0.997 1。

	H	S	D
H	1	0.627 6	0.675 2
S	0.627 6	1	0.997 1
D	0.675 2	0.997 1	1

9.11 下表给出了未经季节调整的衣服和附属品零售季度数据 (1983~1986年)：

服装公司的销售量

(单位：百万美元)

年份	第一季度	第二季度	第三季度	第四季度
1983	4 190	4 927	6 843	6 912
1984	4 521	5 522	5 350	7 204
1985	4 902	5 912	5 927	7 987
1986	5 458	6 359	6 501	8 607

资料来源：Business Statistics，1986，U.S.Department of Commerce，A Supplement of the *Survey of Current Business*，p.37(quarterly averages computed from monthly data).

考虑下面的模型：

$$\text{销售量}_t = B_1 + B_2 D_{2t} + B_3 D_{3t} + B_4 D_{4t} + u_t$$

式中 D_2——1，第二季度；0，其他

D_3——1，第三季度；0，其他

D_4——1，第四季度；0，其他

(a) 对上述回归方程进行估计。

(b) 解释各个系数的含义。

(c) 如何利用估计的回归结果消除季节变动的成分？(选做题)

9.12 利用习题9.11的数据，但现在估计下面的模型：

$$\text{销售量}_t = B_1 D_{1t} + B_2 D_{2t} + B_3 D_{3t} + B_4 D_{4t} + u_t$$

(a) 这个模型与习题9.11的模型有何区别？ (b) 估计这个模型，是否需要略去截距项？换句话说，你是否作通过原点的回归方程？ (c) 比较本题与上题的回归结果，你决定选择哪个模型？为什么？

9.13 参考本章例9.3。如何建立回归模型使得初职年薪的截距和斜率因男女教师的不同而不同？

9.14 参考模型(9-12)。如何解释模型中的非虚拟变量？

9.15 如何对模型(9-13)进行调整使得斜率系数也可能不同？

9.16 考虑下面的模型：

$$Y_i=B_1+B_2D_{2i}+B_3D_{3i}+B_4(D_{2i}D_{3i})+B_5X_i+u_i$$

式中　Y——大学教师的年薪

X——教龄

$D_2=\begin{cases}1, & \text{男教师}\\ 0, & \text{女教师}\end{cases}$　　$D_3=\begin{cases}1, & \text{白种人}\\ 0, & \text{其他}\end{cases}$

(a) $(D_{2i}D_{3i})$表示交互项。它有什么意义？　(c) B_4有什么意义？　(d) 求$E(Y_i \mid D_2=0, D_3=1, X_i)$，并作出解释。

9.17　在模型(9-1)中，假定

$D_i=\begin{cases}1, & \text{大学毕业}\\ 0, & \text{非大学毕业}\end{cases}$

利用表9-1给出的数据,对模型(9-1)回归,将此回归结果与式(9-4)相比较,你能得出什么结论？

9.18　继续上题，但现在假定

$D_i=\begin{cases}2, & \text{大学毕业}\\ 0, & \text{非大学毕业}\end{cases}$

再对模型(9-1)进行估计并对回归结果比较，你得出什么结论？

9.19　表9-8美国1985年第一季度至1991年第一季度的给出了税后公司利润和净利润 (亿美元)的季度数据，

(a) 作(Y)对税后利润(X)的回归，看看两者之间是否相关？　(b) 如果红利支付呈现出季节变动的趋势，将模型引入一个适当的虚拟变量并对其进行估计。在建立模型当中，如何考虑截距和斜率会随季节的不同而变化？　(c) 在不考虑季节因素时，何时作Y对X的回归？

9.20　参考例9.3。利用表9-2提供的数据，并用虚拟变量技术看看大学男女教师的年薪增长率是否不同。如果情况果真如此，你对方程 (9-10)作何评论。

表9-8　美国税后公司利润

年份-季度	红利	税后利润	年份-季度	红利	税后利润
1985-1	87.2	125.2	1988-3	117.5	213.4
1985-2	90.8	124.8	1988-4	121.0	226.0
1985-3	94.1	129.8	1989-1	124.6	221.3
1985-4	97.4	134.2	1989-2	127.1	206.2
1986-1	105.1	109.2	1989-3	129.1	195.7
1986-2	110.7	106.0	1989-4	130.7	203.0
1986-3	112.3	110.0	1990-1	132.3	199.1
1986-4	111.0	119.2	1990-2	132.5	193.7
1987-1	108.0	140.2	1990-3	133.8	196.3
1987-2	105.5	157.9	1990-4	136.2	199.0
1987-3	105.1	169.1	1991-1	137.8	189.7
1987-4	106.3	176.0	1991-2	136.7	182.7
1988-1	109.6	195.5	1991-3	138.1	189.6
1988-2	113.3	207.2	1991-4	138.5	190.3

资料来源：U.S.Department of Commerce，Bureau of Economic Analysis,*Business Statistics,1963-91*,dividend and after-tax profits are in billions of dollars and are obtained from p.A-110.

第三部分

实践中的回归分析

这一部分共包括五章内容(从第10章到第14章)，讨论了线性回归模型在实际应用中存在的若干问题。第二部分建立起的古典线性回归模型(CLRM)虽应用广泛，但却是建立在若干假设基础之上的，而在具体的实践中这些假定并不一定满足。在第三部分中，我们将讨论在具体应用中，一个或若干个CLRM假定不满足时所发生的情形。

第10章讨论多重共线性——两个或多个解释变量相关时所发生的情形。回顾古典线性回归模型的假定之一：解释变量不与其他解释变量完全线性相关。本章指出：只要解释变量不完全线性相关，普通最小二乘估计量(OLS)仍然是最优线性无偏估计量。

第11章讨论异方差——违背古典线性回归模型假定之一，误差项方差为常数时的情形。本章指出：如果违背误差项方差为常数这一假定，则 OLS估计量虽然是无偏的，但却不是有效的。简言之， OLS估计量不是最优线性无偏估计量。本章还指出：怎样通过一些简单变换来消除异方差问题。

第12章讨论自相关——违背古典线性回归模型假定之一，误差项之间相关。与异方差情形相同，如果存在自相关，则 OLS估计量虽然是无偏的，但却不是有效的，所以 OLS估计量也不是最优线性无偏估计量。本章指出如何对数据进行适当的变换使自相关问题最小化。

第13章讨论模型的选择——古典线性回归模型的一些不成文假定，即如何确定所选择的模型是“正确的”模型。本章讨论回归模型的各种错误设定所导致的后果并提出适当的修补措施。

第14章讨论了在实践中广泛应用的一些重要回归模型。主要有： (1)有限最小二乘模型，(2)动态回归模型，(3)应变量为定性变量的回归模型。通过对这些模型的讨论表明：在处理各种实际问题中，线性回归模型是一种非常灵活的工具。本章还讨论了伪回归，即应变量和解释变量之间的关系可能是表面的而非真实的。在本章的最后简短地介绍了在金融领域中常用的随机游走模型。

与前面的章节一样，我们仍用一些数字的和具体的经济实例来阐述在这5章中所出现的一些重要概念。

第10章

多重共线性

第7章中我们列出了古典线性回归模型(CLRM)的假设之一：不存在完全的多重共线性——也即多元回归中的各个解释变量，X_s，之间不存在完全的线性关系。在那一章里，我们直观地解释了完全多重共线性的含义，并说明了在总体多元回归模型(PRF)中为什么要假定不存在完全多重共线性。本章我们主要讨论多重共线性。实际上，在实际中很少遇到完全多重共线性的情况，常常是接近或高度多重共线性(near or very high multicollinearity)的情况，在这些情况下，解释变量是接近线性相关的。重要的是要弄清楚这些相关的变量会对普通最小二乘估计带来什么问题。本章试图回答以下问题：

(1) 多重共线性的性质是什么？
(2) 多重共线性是否真的是一个问题？
(3) 多重共线性的理论后果是什么？
(4) 多重共线性的实际后果是什么？
(5) 在实际中，如何发现多重共线性？
(6) 消除多重共线性的弥补措施有哪些？

表10-1 对widgets的需求

Y (数量)	X_2 (价格/美元)	X_3 (每周收入/美元)	X_4 (每周收益/美元)
49	1	298	297.5
45	2	296	294.9
44	3	294	293.5
39	4	292	292.8
38	5	290	290.2
37	6	288	289.7
34	7	286	285.8
33	8	284	284.6
30	9	282	281.1
29	10	280	278.8

10.1 多重共线性的性质：完全多重共线性的情况

为了回答这几个问题，我们先考虑一个简单的数字例子，我们构造这个例子是为了归纳出

多重共线性的一些基本要点。回顾第6章中对widegts的需求函数。表5-4中的数据已经复制在表10-1的前两列中，并且还加上了一些额外信息。

在表10-1给出了两列收入数据，比方说，X_3是由某一位研究人员所估计的，而X_4则是由另一位研究人员所估计的。为了对两者加以区分，我们把X_3称为收入而把X_4称为收益。

既然对于大多数商品而言，除了价格以外，消费者的收入也是一个重要的因素，我们把需求函数加以扩展，写成：

$$Y_i=A_1+A_2X_{2i}+A_3X_{3i}+u_i \tag{10-1}$$

$$Y_i=B_1+B_2X_{2i}+B_3X_{4i}+u_i \tag{10-2}$$

这两个需求函数的不同之处在于对收入的不同测度。先验地，或者依据理论，我们预计A_2和B_2是负值(为什么？)，而A_3和B_3是正值(为什么？)。[1]

当我们尝试着运用表10-1中的数据来进行回归时，计算机“拒绝”估计回归。[2] 这是怎么回事呢？其实并没有什么。作价格(X_2)和收入(X_3)的关系图，见图10-1。如果我们作X_3对X_2的回归，则得到如下方程：

$$X_{3i}=300-2X_{2i} \qquad R^2(=r^2)=1.00 \tag{10-3}$$

换句话说，收入变量(X_3)与价格变量(X_2)完全线性相关；也即存在完全共线性(或者说是多重共线性)。[3]

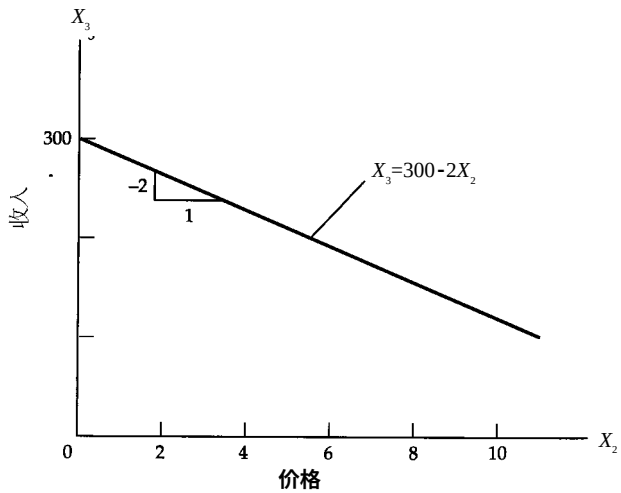


图10-1 收入(X_3)与价格(X_2)散点图

由于存在完全共线性，所以我们不能对方程(10-1)进行回归估计。若把方程(10-3)代入方程(10-1)中，有：

$$\begin{aligned} Y_i&=A_1+A_2X_{2i}+A_3(300-2X_{2i})+u_i \\ &=(A_1+300A_3)+(A_2-2A_3)X_{2i}+u_i \\ &=C_1+C_2X_{2i}+u_i \end{aligned} \tag{10-4}$$

1 按照经济理论，对于大多数经济货物而言，收入系数将会是正值。但对于“劣等”商品而言，这一系数将会是负值。

2 通常地，你将得到一则计算机信息，表明X，或者数据矩阵并未正确加以定义；也就是说，它无法加以转换。在矩阵代数中这样的矩阵被称作是奇异矩阵。简言之，计算机无法进行运算操作。

3 共线性指的是变量之间的存在单一的完全线性关系，但多重共线性是指变量之间存在多种这样的关系。从现在起，我们在用多重共线性来概括这两种情形。通过上下文可以了解是仅有一个亦或是有多个线性关系。

其中，

$$C_1=A_1+300A_3 \tag{10-5}$$

$$C_2=A_2-2A_3 \tag{10-6}$$

难怪我们不能估计方程(10-1)，从式(10-4)可以看出，这并不是多元回归，而是Y对X_2的一个简单的双变量回归。现在，我们虽然可以估计方程(10-4)并获得C_1和C_2的估计值，但根据这些变量我们无法求得原始参数A_1，A_2和A_3的估计值，因为这里只有两个方程(10-5)和(10-6)，但未知参数却有3个。(根据初等代数的知识，我们知道，估计3个变量通常需要3个方程。)

式(10-4)的估计结果已在式(6-48)中给出，为了方便起见，我们将其引用如下：

$$\begin{aligned}\hat{Y}_i&=49.667-2.157\ 6X_{2i}\\ \text{se}&=(0.746)\ (0.120\ 3)\\ t&=(66.538)\ (-17.935)\\ r^2&=0.975\ 7\end{aligned} \tag{10-7}$$

我们可以看到，$C_1=4.9667$，$C_2=-2.157\ 6$。我们无法根据这两个值推出3个未知变量A_1，A_2和A_3的值。[1]

以上讨论的结论是：当解释变量之间存在完全线性相关或者完全多重共线性时，我们不可能获得所有参数的惟一估计值。既然我们不能获得它们的惟一估计值，也就不能根据某一样本做任何统计推论(也即假设检验)。

坦率地说，在完全多重共线性的情况下，不可能对多元回归模型中的某一单个回归系数进行估计和假设检验。这是一个死胡同。当然，正如方程(10-5)(10-6)所示的那样，我们可以求得原始系数的一个线性组合(也即和或者差)的估计值，但却无法获得每个系数的估计值。

10.2 接近或者不完全多重共线性的情形

多重共线性是一个极端情形。在用经济数据进行分析时，两个或多个解释变量之间并不完全线性相关，但却接近完全线性相关。也就是说，共线性的程度可能很高，但并不是完全的共线性。这就是接近，或不完全，或者高度多重共线性的情形。关于“高度”共线性的含义我们将在稍后加以解释。从现在起，我们所说的多重共线性是指不完全多重共线性。正如我们在10.1节所看到的那样，完全多重共线性的情形是一个“死胡同”。

为了弄清楚接近(不完全)多重共线性的含义，我们回到表10-1，但这一次，我们以收益作为收入变量来进行回归。回归的结果如下：

$$\begin{aligned}\hat{Y}_i&=14.537-2.797\ 5X_{2i}-0.319\ 1X_{4i}\\ \text{se}&=(120.06)(0.8122)(0.400\ 3)\\ t&=(1.210\ 7)(-3.444\ 4)(-0.797\ 1)\\ R^2&=0.977\ 8\end{aligned} \tag{10-8}$$

这些结果是比较有意思的，原因如下：

(1) 尽管不能估计回归方程(10-1)，但我们能够估计回归方程(10-2)，虽说两种收入变量之间的差别是相当小的，这一点从表10-1的最后两列能够直观地看出来。[2]

1 当然，如果A_1，A_2和A_3的任何一个值是固定的，则另外两个A的值是可以通过所估计的C来求得的。但是这些值不是惟一的，因为A是从3个中任取的一个。重申一遍，根据两个已知变量(两个C)无法求得3个未知变量(3个A)的值。

2 我们可以揭开“庐山真面目了”。在表10-1第4列中的收益数据是通过下列关系而得来的：$X_{4i}=X_{3i}+u_i$其中u是从随机数表中获得的随机数。这10个随机数u的值如下：-0.5，-1.1，-0.5，0.8，0.2，1.7，-0.2，0.6，-0.9和-1.2。

(2) 与预期相同，方程(10-7)和(10-8)中的价格系数都是负的，并且两者之间的数值差异并不大。每一价格系数都显著地不为零 (为什么？)。但是，相对而言，方程 (10-7)中的价格系数的$|t|$值略大于方程(10-8)中相应的$|t|$值。或者说，方程(10-7)的价格系数的标准差 (se)略小于方程(10-8)中系数的标准差。

(3) 方程(10-7)的R^2值为0.975 7，而方程(10-8)的R^2值为0.977 8，增加量只有0.002 1，这一增加量并未表现出是一个大的增加量。可以证明，R^2值的这一增加量是统计不显著的。[1]

(4) 收入(收益)变量的系数是统计不显著的，但更为重要的是，它的符号是错误的：对于大多数商品来说，收入对于商品的需求量具有正向影响，除了劣等商品以外。

(5) 尽管收入变量是不显著的，但若检验假设：$B_2=B_3=0$(也即假设$R^2=0$)，则根据方程(7-50)给出的F检验，很容易拒绝该假设。换句话说，价格和收益集体地或联合地对商品的需求有显著影响。

如何解释这些“奇怪”的结果呢？作价格X_2与收入X_4的关系图(参见图10-2)。从图中可以看到，与图10-1不同，尽管价格和收益并不完全线性相关，但两个变量之间却存在高度的依存关系。这一点可以从下面的回归方程中更清楚地看出：

$$\begin{aligned} X_{4i} &= 299.92 - 2.005\,5X_{2i} + e_i \\ \text{se} &= (0.674\,8)(0.108\,8) \\ t &= (444.44)(-18.44) \\ r^2 &= 0.977\,0 \end{aligned} \tag{10-9}$$

回归结果表明：价格与收益高度相关；相关系数为- 0.9884(即r^2的负平方根)。也即接近完全线性相关，或是接近完全多重共线性。如果相关系数为- 1，如方程(10-3)，则是完全多重共线性的情形。特别需要注意的是：在方程 (10-3)中，我们没有加上e_i这一项，因为X_{2i}和X_{3i}之间线性关系是完全的，但在方程 (10-9)中却加上e_i这一项以表明X_{4i}和X_{2i}之间线性关系是不完全的。

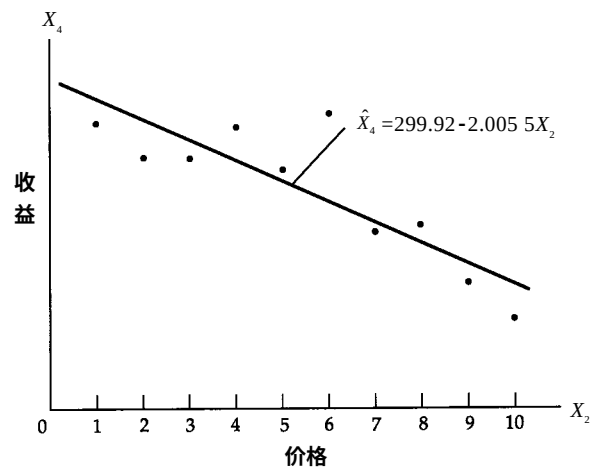


图10-2 收益(X_4)和价格(X_2)关系

在结束这一节的讨论时，需要指出的是：在只有两个解释变量的情形下，相关系数 r可用作共线性程度的测度。但当解释变量多于两个时，相关系数则不能充分测度共线性。

1 这是解释变量对拟合方程的“边际”或者“增量”贡献的问题。对这一问题的详细讨论见第14章第14.1节。此外，还可参阅Damodar N Gujarati，*Basic Econometrics*, 3d ed., McGraw-Hill, New York, 1995, pp.250-253。

10.3 多重共线性的理论后果

到目前为止，我们已经讨论了完全和不完全多重共线性的性质。下面我们讨论多重共线性引起的后果。但需记住：从现在起我们仅考虑不完全多重共线性的情形，因为讨论完全多重共线性有些不切实际。

我们知道，在古典线性回归模型 (CLRM)的假定下，普通最小二乘 (OLS)估计量是最优线性无偏估计量 (BLUE)：在所有线性无偏估计量中，普通最小二乘估计量具有最小方差性。引人注意的是，即使共线性是不完全的，普通最小二乘估计量仍然是最优线性无偏估计量，即使多元回归方程的一个或多个偏回归系数是统计不显著的。在方程 (10-8)中，价格系数是统计显著的，但收入系数却是统计不显著的。但在方程 (10-8)中所用的普通最小二乘法仍保持其最优线性无偏估计量的性质。[1]既然如此，我们为什么还“大惊小怪”地讨论多重共线性呢？这主要有如下几个原因：

(1) 的确如此，即使是在接近共线性的情形下，普通最小二乘法估计量仍然是无偏的。但要记住的是，无偏性是一个重复抽样的性质。也就是说，保持 X不变，如果得到一些样本并用普通最小二乘法计算这些样本估计量，则其平均值收敛于估计量的真实值。但这并不是某个样本估计值的性质。在现实中，我们无法得到大量的重复样本。

(2) 接近共线性也并未破坏普通最小二乘估计量的最小方差性：在所有线性无偏估计量中，普通最小二乘法估计量的方差最小。然而这并不意味着在任何给定的样本中普通最小二乘法估计量的方差会很小 (与估计量的值本身相比而言)，这一点在方程 (10-8)中表现得非常清楚。的确，收入系数估计量是最优线性无偏估计量，但与估计值相比，样本的方差如此之大以至于计算的 t值(在零假设：真实收入系数为零)仅有 - 0.797 1。这将使我们接受假设，也即收入对于 widgets的需求量没有任何影响。简言之，最小方差并不意味着方差值也较小。

(3) 即使变量 X与总体(也即PRF)不线性相关，但却可能与某一样本线性相关，例如表 10-1。从这个意义上说，多重共线性本质上是一个样本 (回归)现象。在假定概率分布函数时，我们认为模型中所有变量 X_s都对应变量 Y有独立的或单独的影响。但也有可能出现这种情形，即在用某个样本估计总体 PRF时，部分或者所有变量 X_s高度共线以至于无法区分它们各自对应变量 Y 的影响。也就是说，虽然理论表明所有的 X_s都是重要的，但是样本却令人失望。之所以会这样是因为大多数经济数据都不是通过试验获得。有些数据，比如国民生产总值、价格、失业率、利润、红利等等，是以其实际发生值为依据，而并非试验得到。如果这些数据能够通过试验获得，那么从一开始我们就不允许共线性的存在。既然如此，如果两个或者多个解释变量之间存在接近共线性，那么我们也“巧妇难为无米之炊”。[2]

10.4 多重共线性的实际后果

在接近或者高度多重共线性的情形下，正如对 widget需求的一例那样，我们可能遇到如下一个或者多个后果：

(1) 普通最小二乘法估计量的方差和标准差较大。这一点可以从回归方程 (10-7)和(10-8)中清楚地看到。我们前面讨论过，由于在价格 (X_2)和收益(X_4)之间存在高度共线性，当方程 (10-8)同时包括这两个变量时，与方程 (10-7)相比，价格变量系数的标准差显著地增加了。我们知道，

1 因为不完全多重共线性本身并未违背我们在第 7章中所罗列的假设，普通最小二乘法估计量仍具有最优线性无偏估计量这一性质。

2 J.Johnston，*Econometric Methods*, 2d ed., McGraw-Hill, New York，1972，p.164.

如果估计量的标准差增加了，则对估计量真实值的估计就变难了。也就是，普通最小二乘法估计量的精确度(precision)就下降了。

(2) 置信区间变宽。由于标准差较大，所以总体参数的置信区间也就变大了。

(3) t值不显著。在检验假设：真实的 $B_3=0$(回归方程(10-8))，我们计算出t值，$b_3/\text{se}(b_3)$，并与t临界值作比较。但是，如前所述，当存在高度共线时，由于估计的标准差急剧地增加，因而使得t值变小。因此，在这重情况下，我们会很自然地接受零假设，即真实的总体系数为零。因而，在回归方程(10-8)中，由于t值为-0.7971，我们也许会断然得出结论：在widgets一例中，收入对于需求的数量没有任何影响。

(4) R^2值较高，但t值则并不都显著。从回归方程(10-8)可以清楚地看出：R^2值很高，约为0.98，但只有价格变量的t值是显著的。根据F(检验)值，我们能够拒绝假设：价格和收益变量同时对widgets的需求没有任何影响。

(5) 普通最小二乘估计量及其标准差对数据的微小变化非常敏感；也就是说，它们趋于不稳定。让我们回到表10-1。假设我们稍微改变一下收益变量X的数据。第一，第五和第十个观察值现在分别为295，287和274，其他数据不变。这一变化的结果如下：

$$\begin{aligned}\hat{Y}_i &= 100.56 - 2.516\,4X_{2i} - 0.169\,95X_{4i}\\ \text{se} &= (48.030)\ (0.359\,06)\ (0.160\,4)\\ t &= (2.093\,6)\ (-7.008\,3)\ (-1.059\,7)\\ R^2 &= 0.979\,1\end{aligned} \tag{10-10}$$

比较方程(10-8)和方程(10-10)，我们注意到，数据的一个较小变化却导致了回归方程的较大变化。相对而言，方程(10-10)中的标准差变小了，从而导致t值的绝对值增大了；与以前相比，收入变量的系数也变小了。

为什么会有这一变化呢？在方程(10-8)中，X_2和X_4之间的相关系数为-0.988 4，而在方程(10-10)中相关系数为-0.943 1。换句话说，从方程(10-8)到方程(10-10)，X_2和X_4之间的共线性程度降低了。尽管相关系数降低的程度不大，但回归结果的变化却是值得注意的。而这正是接近完全共线性引起的后果。

(6) 回归系数符号有误。如回归方程(10-8)和(10-10)所示，收益变量的符号是“错误”的，因为按照经济理论，我们知道对于大多数商品而言，收入效应是正的。当然，对于劣等品而言，这并不是“错误”的符号。

(7) 难以衡量各个解释变量对回归平方和(ESS)或者R^2的贡献。我们仍用widgets一例来说明。在方程(10-7)中，我们用仅价格(X_2)来拟合需求量(Y)，得到R^2为0.975 7。而在方程(10-8)中，我们用价格和收益来拟合Y，得到R^2为0.977 8。现在如果我们仅用X_4来拟合Y，则我们可以得到如下结果：

$$\begin{aligned}\hat{Y}_i &= -263.74 + 1.043\,8X_{4i}\\ \text{se} &= (26.929)\ (0.093\,2)\\ t &= (-9.794)\ (11.200)\\ R^2 &= 0.940\,0\end{aligned} \tag{10-11}$$

你会发现，仅收益X_4就解释了需求量变化的94%。此外，收益系数不仅在统计上是显著的，而且其符号也是正的，与理论预期一致。

如前所述，在多元回归方程(10-8)中，R^2值为0.977 8。哪一部分归于X_2，哪一部分归于X_4呢？我们并不能精确地加以区分，因为这两个变量高度共线性以至于当一个变量变化时，另一个变量也几乎是自动地随之变化，回归方程(10-9)清楚地表明了这一点。因此，在高度共线性情况下，很难衡量每一个解释变量对总体R^2的贡献。

问题是，我们此前所阐述的多重共线性的后果能否严格地确定下来呢？当然可以！这里我们略去证明，感兴趣的同学可以参阅有关参考书。[1]

10.5 多重共线性的测定

如前节所述，当多重共线性程度较高时，利用普通最小二乘法求参数估计值会产生一系列的严重后果。那么，如何解决多重共线性问题呢？在解决这一问题之前，我们首先需要确定是否存在共线性问题。简言之，我们关心的是多重共线性存在与否以及多重共线性的程度。我们知道多重共线性是一个样本特性(sample specific)，它是一个样本现象。因此，下面的一些警告是必要的：[2]

(1) 多重共线性是一个程度问题而不是存在与否的问题。

(2) 由于多重共线性是在假定解释变量是非随机的条件下出现的问题，因而它是样本的特征，而不是总体的特征。

因此，我们不仅可以“检测多重共线性”，而且可以测度任何给定样本的多重共线性程度。

需要补充的是：我们并没有多重共线性的单一测度方法，因为对于非试验数据，我们无法确定其共线性的性质与程度。我们所具有是一些经验法则，或者说是在具体应用中能够给我们提供一些有关多重共线性存在与否的线索。比如：

(1) R^2较高但t值显著的不多。如在前面讲到的那样，这是多重共线性的“经典”(classic)特征。如果R^2较高，比如说超过了0.8，在大多数情况下F检验将会拒绝零假设，即部分斜率系数联合(jointly)或同时(simultaneously)为零。但各自的t检验表明，没有或几乎没有部分斜率系数是统计显著不为零的。式(10-8)回归结果就充分地证实了这一点。

(2) 解释变量两两高度相关(high pairwise correlations)。例如，如果在多元回归方程包括6个解释变量，运用式(6-44)计算这些变量两两之间的相关系数，如果有些相关系数很高，比方说超过0.8，则可能存在较为严重的共线性。遗憾的是，这一标准并不十分可靠，因为解释变量两两相关系数可能较低(表明无高度共线性)，但是却有可能存在共线性，因为t值中很少是统计显著的。[3]

(3) 检验解释变量相互之间的样本相关系数。假设我们有三个解释变量，X_2，X_3和X_4。我们以r_{23}，r_{24}和r_{34}来分别表示X_2与X_3，X_2与X_4以及X_3与X_4之间的两两相关系数。假设r_{23}=0.90，表明X_2和X_3之间高度共线性。现在来考虑相关系数，$r_{23.4}$，我们称之为偏相关系数(partial correlation coefficient)，它是在变量X_4为常数的条件下，X_2和X_3之间的相关系数(这一概念有些类似于我们在第7章所讨论过的偏回归系数)。假设$r_{23.4}$=0.43，也就是说，在变量X_4保持不变的条件下，X_2和X_3之间的相关系数仅为0.43，但是若不考虑X_4的影响，这一相关系数为0.90。于是，根据偏相关系数，我们不能判断X_2与X_3之间的共线性程度一定很高。

因此，在存在多个解释变量的情形下，依赖简单的两两相关系数来判断多重共线性可能会“误入歧途”。不幸的是，用偏相关系数代替简单的两两相关系数并未提供一个检验多重共线性存在与否的确切依据。而后者仅仅是检验多重共线性性质的另一手段。[4]

---

1 证明可参阅Damodar N Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York,1995，pp.327-332。

2 Jan Kmenta, *Elements of Econometrics*,2d ed.,Macmillan，1986，pp.431。

3 更详细的讨论参阅Damodar N Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York,1995，pp.335-336。

4 参阅Damodar N Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York,1995,pp.335-336。

(4) 从属(subsidiary)或者辅助(auxiliary)回归。既然多重共线性是由于一个或多个解释变量是其他解释变量的线性(或接近线性)组合，那么检验模型中哪个变量与其他变量高度共线性的方法就是作每个变量对其他剩余变量(remaining)的回归并计算相应的R^2值。其中的每一个回归都被称作是从属或者辅助回归，从属于Y对所有变量的回归。

例如，考虑Y对X_2，X_3，X_4，X_5，X_6、X_7——这6个解释变量的回归。如果回归结果表明存在多重共线性，比方说，R^2值很高，但解释变量的系数很少是统计显著的，那么，究其根源，即找出哪些变量可能是其它变量的线性(或接近线性)组合。具体步骤如下：

(1) 作X_2对其他剩余解释变量的回归，并求样本决定系数，记为R_2^2。

(2) 作X_2对其他剩余解释变量的回归，并求样本决定系数，记为R_3^2。

对模型中剩余的解释变量继续以上步骤。在这个例子中，总共有6个这样的辅助回归。

我们如何判断哪些解释变量是共线性的呢？估计的R_i^2值介于0和1之间(为什么？)。如果某个解释变量不是其他变量的线性组合，则该回归方程的R_i^2值不会显著不为零。根据方程(7-50)，我们知道应该如何去检验假设：某个样本决定系数显著为零。

假定我们想要检验假设：$R_2^2=0$，也即X_2与剩余5个解释变量并不存在共线性。根据式(7-50)，即：

$$F=\frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

其中n是观察值的个数，k是包括截距在内的解释变量的个数。具体说明如下。

在这个例子中，假设有一个容量为50的随机样本，对每个解释变量作剩余变量的回归。各辅助回归的R^2值如下：

$R_2^2=0.90$ (X_2对于其他变量的回归)
$R_3^2=0.18$ (X_3对于其他变量的回归)
$R_4^2=0.36$ (X_4对于其他变量的回归)
$R_5^2=0.86$ (X_5对于其他变量的回归)
$R_6^2=0.09$ (X_6对于其他变量的回归)
$R_7^2=0.24$ (X_7对于其他变量的回归)

表10-2给出F检验的结果。

表10-2 检验R^2的显著性(方程7-50)

R^2值	F值	F值是否显著？
0.90	79.20	是①
0.18	1.93	否
0.36	4.95	是①
0.86	54.06	是①
0.09	0.87	否
0.24	2.87	是②

① 显著水平为1%
② 显著水平为5%
本例中n=50，k=6

如表所示，变量X_2，X_4，X_5和X_7看来与其他变量共线性，尽管共线性程度(用R^2度量)差别很大。由此得出的结论是：“看似”较低的R^2，比如0.36，却可能是统计显著不为零。我们将在10.7节中给出辅助回归的一个具体经济实例。

辅助回归技术的一个缺陷是它的计算较繁琐。如果一个回归方程包含若干个解释变量，则我们不得不计算好几个辅助回归方程，因此，这种方法没有实用价值。但需要指出的是：现在已有许多统计软件可以用来计算辅助回归方程。

(5) 方差膨胀因素(the variance inflation factor)。即使模型并未包括太多的解释变量，从各个辅助回归方程中获得的R^2值也未必可以用于诊断共线性。我们可以从第7章中讨论过的一个三变量回归方程中更清楚地看到。方程(7-25)和(7-27)给出了计算两个部分斜率系数b_2和b_3方差的公式。通过简单的代数变换，这些方差公式可以写为：

$$\text{var}(b_2)=\frac{\sigma^2}{\sum x_{2i}^2(1-R_2^2)} \tag{10-12}$$

$$=\frac{\sigma^2}{\sum x_{2i}^2}\text{VIF}$$

$$\text{var}(b_3) = \frac{\sigma^2}{\sum x_{3i}^2 (1 - R_2^2)} = \frac{\sigma^2}{\sum x_{3i}^2} \text{VIF} \tag{10-13}$$

(证明见习题10.21)在这些公式中，R_2^2 是X_2与X_3之间(辅助)回归方程的样本决定系数 (注意，X_2与X_3之间的R^2值等于X_3与X_2之间的R^2值)

在上述公式中，

$$\text{VIF} = \frac{1}{(1 - R_2^2)} \tag{10-14}$$

方程(10-14)右边的表达式形象地称作方差膨胀因素 (VIF)，因为随着R^2的增加，b_2和b_3的方差(标准差)也增加了或者说是膨胀了 (你看出这一点了吗？)。更极端地，如果样本决定系数为1(也即完全多重共线性)，则这些方差和标准差将会没有意义 (为什么？)。当然，如果R^2为零，那么就不存在共线性， VIF的值为1(为什么？)。我们也就不必担心由于方差 (标准差)较大而带来的问题。

现在一个重要的问题是，假设在辅助回归方程中，R_i^2 值很高(但小于1)，表明存在较高程度的共线性。但是从方程(10-12)，(10-13)和(10-14)可以清楚地看到，比如说，b_2的方差不仅仅取决于VIF，而且还取决于u_i的方差σ^2和X_2的方差 $\sum x_{2i}^2$ 。因此，以下情形是很有可能的：R_i^2 值很高，比如说是0.91，但是σ^2较低或者 $\sum x_{2i}^2$ 较高，或者是两种情况同时出现，以至于 b_2的方差较低，t值较高。换句话说，较高的R^2可能被一个较低的σ^2值或者较高的 $\sum x_{2i}^2$ 值所抵消。当然，高和低仅是相对而言。

所有这些都表明，辅助回归方程中的 R_i^2 可能只是多重共线性的一个表面指示器 (surface indicator)。如上所述，它并不一定扩大估计量的标准差。更正规地表述为，“ 较高的 R_i^2 既不是较高标准差的必要条件也不是充分条件。多重共线性本身并不必然导致较高的标准差。”[1]

从上面讨论的各种多重共线性的检验方法中，我们能得出哪些结论呢？检验多重共线性有许多种不同的方法，但却没有一种检验方法能够使我们彻底解决多重共线性问题。记住一点：多重共线性是一个程度的问题，它是与样本相关的一种现象。有些时候，可以“ 容易 ”地检验出多重共线性，但有些时候，我们却必须综合运用上面讨论的各种手段来诊断这一问题的严重程度。总之，没有一个简单的方法解决这个问题。

对多重共线性检验的研究还将继续。现在，已有一些新方法，比如病态指数(condition index)等。对这些方法的讨论已超出本书研究的范围，有兴趣的同学可参阅有关参考书。[2]

10.6 多重共线性必定不好吗

在我们讨论多重共线性补救措施之前，需要回答一个重要的问题：多重共线性必然是一个“ 恶魔 ”吗？答案取决于研究的目的。如果研究是为了用模型来预测解释变量的未来均值，则多重共线性本身未必是一件坏事。回到对 widgets的需求方程(10-8)中，尽管收益变量不是统计

1 G.S.Maddala，*Introduction to Econometrics*，Macmillan，New York，1988，p.266。但是，Maddala也曾说，“ 如果R_i^2较低，则我们的境遇会好一些。”

2 关于条件指数的简单讨论，参见 Damodar N Gujarati，*Basic Econometrics*，3d ed.，，McGraw-Hill，New York，1995，pp.338.

显著的，但总体R^2值，0.977 8，还是略高于略去了收益变量的方程(10-7)中的R^2值。因此，出于预测的目的，方程(10-8)略好于方程(10-7)。通常，预测人员都是根据解释能力(用R^2度量)来选择模型的。这是否是一个好的标准呢？如果我们假定表10-1中价格和收益数据之间所观察到共线性一直保持下去，则它也许是一个好的判定标准。在方程(10-9)中，我们已经证明了X_4和X_2(收益和价格)是如何相关的。如果同样的关系一直继续下去，则方程(10-8)可以用于预测。但那仅仅是如果。如果在另一样本中，两个变量的共线性程度并没有那么高，显然，依据方程(10-8)的预测就没有什么价值。

另一方面，如果研究不仅仅是为了预测，而且还要可靠地估计所选模型的各个参数，则严重的共线性将是一件“坏事”，因为它将导致估计量的标准差增大。然而，正如我们前面指出的那样，如果是为了比较准确地估计一组系数(例如，两个系数的和或者差)，那么，即使存在多重共线性，也能够达到目的。在这种情形下，多重共线性或许不是什么问题。因而，在方程(10-7)中，斜率-2.157 6是(A_2-2A_3)的估计值(我们可以用普通最小二乘法准确地估计，虽说不能单独地估计出A_2和A_3的值)。

也存在着一些“令人愉快的”情形：尽管存在高度共线性，但在常用的显著水平，如5%的显著水平下，根据t检验，估计的R^2和大多数单个回归系数都是统计显著的。正如Johnston所指出的那样：

这倒是有可能产生的，如果各个系数正好在数值上远超过其真实值，那么尽管标准差扩大了，效应依然显示出来而且(或者)由于真实值本身很大，以至于即使估计值过低，仍然表现出是统计显著的。

在继续下面的讨论之前，让我们花点时间来看一个具体的经济例子以说明上述讨论的几个要点。

10.7 一个扩充例子：1960至1982年期间美国的鸡肉需求

表10-3给出了美国1960至1982年期间有关数据：平均每人鸡肉消费量(Y)，每人实际(即通货膨胀调整后的)可支配收入(X_2)，鸡肉的实际零售价格(X_3)，猪肉实际零售价格(X_4)，牛肉实际零售价格(X_5)

表10-3 鸡肉需求的原始数据[1]

Y②	$X_2$③	$X_3$④	$X_4$⑤	$X_5$⑥
27.800 00	397.500 0	42.200 00	50.7000 0	78.300 00
29.900 00	413.300 0	38.100 00	52.0000 0	79.200 00
29.800 00	439.200 0	40.300 00	54.0000 0	79.200 00
30.800 00	459.700 0	39.500 00	55.3000 0	79.200 00
31.200 00	492.900 0	37.300 00	54.7000 0	77.40000
33.300 00	528.600 0	38.100 00	63.7000 0	80.200 00
35.600 00	560.300 0	39.300 00	69.8000 0	80.400 00
36.400 00	624.600 0	37.800 00	65.9000 0	83.900 00
36.700 00	666.400 0	38.400 00	64.5000 0	85.500 00
38.400 00	717.800 0	40.100 00	70.0000 0	93.700 00
40.400 00	768.200 0	38.600 00	73.2000 0	106.10 00
40.300 00	843.300 0	39.800 00	67.8000 0	104.80 00
41.800 00	911.600 0	39.700 00	79.1000 0	114.00 00
40.400 00	931.100 0	52.100 00	95.4000 0	124.10 00

1 J.Johnston，*Econometric Methods*，3d ed., McGraw-Hill，New York，1984，p.249。

(续)

Y②	$X_2$③	$X_3$④	$X_4$⑤	$X_5$⑥
40.700 00	102 1.500	48.900 00	94.2000 0	127.60 00
40.100 00	116 5.900	58.300 00	123.500 0	142.90 00
42.700 00	134 9.600	57.900 00	129.900 0	143.60 00
44.100 00	144 9.400	56.500 00	117.600 0	139.20 00
46.700 00	157 5.500	63.700 00	130.900 0	165.50 00
50.600 00	175 9.100	61.600 00	129.800 0	203.30 00
50.100 00	199 4.200	58.900 00	128.000 0	219.60 00
51.700 00	225 8.100	66.400 00	141.000 0	221.60 00
52.900 00	247 8.700	70.400 00	168.200 0	232.60 00

① 实际价格是由名义价格除以消费者食品价格指数而获得的。

② Y——平均每人鸡肉消费量(磅)*

③ X_2——实际可支配收入(美元)

④ X_3——每磅鸡肉实际零售价格(美分)

⑤ X_4——每磅猪肉实际零售价格(美分)

⑥ X_5——每磅牛肉实际零售价格(美分)

资料来源：Y来自城市数据库(CITIBASE)，X_2到X_5来自美国农业部。这些数据均由作者的学生罗伯特 J.费雪(Robert J. Fisher)收集。

从理论上说，商品的需求通常是消费者实际收入、该商品实际价格以及竞争(competing)商品或互补(complementary)商品的实际价格的函数。估计的需求函数如下：应变量(Y)是平均每人鸡肉消费量的自然对数。

解释变量	系 数	标准差se	t值	ρ值	
常数	2.189 8	0.155 7	14.06 3	0.000 0	
$\ln X_2$	0.342 6	0.083 3	4.114 0	0.000 3	
$\ln X_3$	-0.504 6	0.110 9	-4.550	0.000 1	(10-15)
$\ln X_4$	0.148 6	0.099 7	1.490 3	0.767	
$\ln X_5$	0.091 1	0.100 7	0.904 6	0.187 8	
	R^2=0.982 3		$\bar{R}^2$=0.978 4		

由于我们拟合的是对数-线性需求函数，因此，所有的系数都是Y对相应X变量的偏弹性(partial elasticities)。因而，需求的收入弹性约为0.34，需求的自价格弹性约为-0.51，需求(猪肉)的交叉弹性约为0.15，需求(牛肉)的交叉弹性约为0.09。

回归结果表明：需求的收入和自价格弹性各自都是统计显著的，但两个交叉弹性则不是显著的。顺便指出的是：由于收入弹性小于1，所以鸡肉并非是奢侈消费品。鸡肉的需求对于其自身价格是缺乏弹性的，因为其弹性系数的绝对值小于1。

尽管两个交叉价格弹性是正的，表明其它两种肉类与鸡肉相互竞争的，但是它们却不是统计显著的。因而，看来鸡肉的需求并不受猪肉和牛肉价格变化的影响。但这也许是一个草率的结论，因为我们不得不考虑多重共线性的可能性。因此，让我们来考虑在10.5节中所讨论的多重共线性的检验方法。

鸡肉需求函数共线性的检验[方程(10-15)]

相关矩阵(correlation matrix)。表10-4给出了4个解释变量(对数形式)之间的两两相关系数。

* 1磅=0.453 592 37kg。

从表中可以看出，解释变量之间的两两相关系数都很高，实际收入与牛肉价格之间的两两相关系数约为0.98；猪肉价格与牛肉价格之间的两两相关系数约为 0.95；实际收入与鸡肉价格之间的两两相关系数为0.91，等等。尽管两两相关系数很高，并不表明需求函数中一定存在着共线性，只是有可能性存在。

辅助回归。当我们对每个解释变量与其他剩余解释变量进行回归时，发现存在共线性问题。这一点可以从表10-5中所给出的结果看出来。如该表所示，表中所有的 R^2值都超过了0.94；根据式(7-50)F检验，所有这些 R^2都是统计显著的(参见习题10.24)，表明回归方程(10-15)中的每个解释变量都与其它解释变量高度共线性。

表10-4 方程(10-15)中解释变量两两相关系数

	$\ln X_2$	$\ln X_3$	$\ln X_4$	$\ln X_5$
$\ln X_2$	1	0.907 2	0.972 5	0.979 0
$\ln X_3$	0.907 2	1	0.946 8	0.933 1
$\ln X_4$	0.972 5	0.946 8	1	0.954 3
$\ln X_5$	0.979 0	0.933 1	0.954 3	1

注：相关矩阵是对称的。因而，$\ln X_4$和$\ln X_3$之间的相关系数等于$\ln X_3$和$\ln X_4$之间的相关系数。

表10-5 辅助回归

$$\ln \hat{X}_2 = 0.946\,0 - 0.832\,4\ln X_3 + 0.948\,3\ln X_4 + 1.017\,6\ln X_5$$
$$t=(2.556\,4)\quad(-3.490\,3)\qquad(5.659\,0)\qquad(6.784\,7)$$
$$R^2=0.984\,6$$

$$\ln \hat{X}_3 = 1.233\,2 - 0.469\,2\ln X_2 + 0.669\,4\ln X_4 + 0.595\,5\ln X_5$$
$$t=(8.005\,3)\quad(-3.490\,3)\quad(4.865\,2)\quad(3.784\,8)$$
$$R^2=0.942\,8$$

$$\ln \hat{X}_4 = -1.012\,7 + 0.661\,8\ln X_2 + 0.828\,6\ln X_3 - 0.469\,5\ln X_5$$
$$t=(-3.710\,7)\quad(5.659\,0)\quad(4.865\,2)\quad(-2.287\,9)$$
$$R^2=0.975\,9$$

$$\ln \hat{X}_5 = -0.705\,7 + 0.695\,6\ln X_2 + 0.721\,9\ln X_3 - 0.459\,8\ln X_5$$
$$t=(-2.236\,2)\quad(6.784\,7)\quad(3.784\,8)\quad(-2.287\,0)$$
$$R^2=0.976\,4$$

因此，很有可能在方程(10-15)中我们没有发现猪肉和牛肉价格变量的系数在统计上是显著的。但是，这与我们先前讨论过的高度多重共线性的理论后果是一致的。有意思的是，尽管存在高度共线性，但是实际收入和自身价格变量的系数却是统计上显著的。这也许可以从约翰斯顿(Johnston)所提及的事实(参见脚注15)中得到很好地解释。

这一例子表明，在存在高度共线性的情形下，判断一个解释变量是否独自显著时，我们应特别小心。我们将把这一例子带入到下一节，在该节中我们将考虑多重共线性的补救措施。

10.8 如何对付多重共线性：补救措施

假定根据10.5节中讨论的一种或者多种检验方法，发现存在着多重共线性。我们采取什么方法减少共线性的严重程度而不是彻底地消除它？不幸的是，与共线性检验一样，并不存在万无一失的补救措施，只有一些经验的法则。之所以如此是因为多重共线性是某一特定样本的特征而并不是总体的特征。此外，尽管存在着接近共线性，普通最小二乘估计量仍然保持最优线性无偏估计量的性质。的确也存在这样的情况：一个或者多个回归系数是统计不显著的，或者它们中的一些系数的符号是错误的。如果研究人员专心于削弱共线性的严重程度，则可尝试用以下的一种或者多种方法。需要注意的是，如果特定的样本是“性状不佳”的，则没有太多的

补救措施。本着这一告诫，让我们来考虑在经济计量学文献中所讨论过的各种补救措施。

10.8.1 从模型中删掉不重要的解释变量

对待严重的多重共线性问题，最简单的解决方法似乎就是删掉一个或者多个共线性变量。因而，在对鸡肉的需求函数(10-15)中，既然三个价格变量高度相关，为什么不从模型中删掉猪肉价格和牛肉价格变量呢？

但是这一补救措施也许比“疾病”(多重共线性)本身还糟糕。当我们构建一个经济模型，如回归方程(10-15)，总是以一定的经济理论为基础。在我们的这个例子中，依据经济理论，我们预期所有这三个价格都对鸡肉的需求有着一定的影响，因为这三种肉产品在某种程度上是竞争性产品。因此，从经济学的角度上说，回归方程(10-15)是一个恰当的需求函数。遗憾的是，根据表10-3所给的特定样本得出的回归结果，我们并不能确定猪肉和牛肉价格各自对鸡肉需求数量的影响。但是从模型中删掉这些变量将会导致所谓的“模型设定误差”(specification error)，我们将在第13章中详细讨论。随后我们将会看到，如果为了消除共线性问题而从模型中删除一个解释变量，并对缺少这一变量的模型估计，则所估计的参数可能是有偏的。为了更清楚地了解，我们给出删除猪肉和牛肉价格变量之后的回归结果：

$$\ln \hat{Y} = 2.032\ 8 + 0.451\ 5 \ln X_2 - 0.372\ 2 \ln X_3 \tag{10-16}$$

$$t = (17.497)\ (18.284)\ (-5.864\ 7)$$

$$R^2 = 0.980\ 1; \qquad \bar{R}^2 = 0.978\ 1$$

回归结果表明：与方程(10-15)相比，收入弹性增大了，但是价格弹性的绝对值却下降了。换句话说，简化了的模型的估计系数是有偏的(第13章对此有更详细的阐述)。

上述讨论表明：存在着一个两难的问题(trade-off)：为了削弱共线性的严重程度，我们得到的系数估计值可能是有偏的。建议不要仅仅因为共线性很严重就从一个经济上可行的模型中删除变量。所选的模型是否符合经济理论当然是一个重要的问题，我们将在第13章中详细讨论。顺便提及的是：回归方程(10-15)中猪肉价格系数的t值大于1。因此，根据第7章的讨论(见7.11节)，如果我们从模型中删除这一变量，调整后的R^2将会下降，实际情况也是这样。

10.8.2 获取额外的数据或者新的样本

既然多重共线性是一个样本特征，那么在包括同样变量的另一样本中，共线性也许不像第一个样本那样高。关键是我们是否能够得到另一个样本，因为收集数据的费用很高。

有的时候仅仅通过获得额外的数据——增加样本的容量——就能够削减共线性的程度。这一点可以从式(10-12)和式(10-13)中很容易地看出来。例如，在公式

$$\text{var}(b_3) = \frac{\sigma^2}{\sum x_{3i}^2 (1 - R_2^2)} \tag{10-12}$$

中，对于给定的σ^2和R^2，如果X_3样本容量增加了，则$\sum x_{3i}^2$通常将会增加(为什么？)，结果b_3的方差将会减小，标准差也随之减小。

我们考虑下面的消费支出(Y)对于收入(X_2)和财富(X_3)的回归方程：(这里有10个观察值[1])：

$$\hat{Y}_i = 24.337 + 0.871\ 64 X_{2i} - 0.034\ 9 X_{3i}$$

$$\text{se} = (6.280\ 1)\ (0.314\ 38)\ (0.030\ 1) \tag{10-17}$$

$$t = (3.875)\ (2.772\ 6)\ (-1.159\ 5)$$

1 我非常感谢阿尔伯特·扎克(Albert Zucker)提供了式(10-17)和式(10-18)的回归结果。

$$R^2=0.968\ 2$$

回归结果表明：在5%的显著水平下，财富系数不是统计显著的。

但是当样本容量增加到40个观察值时，我们得到如下回归结果：

$$\hat{Y}_i=2.090\ 7+0.729\ 9X_{2i}+0.060\ 5X_{3i}$$
$$t=(0.871\ 3)\ (6.001\ 4)\ (2.064\ 1) \tag{10-18}$$
$$R^2=0.967\ 2$$

现在，在5%的显著水平下，财富系数是统计显著的。

当然，与获得一个新样本一样，出于成本和其他一些因素的考虑，获取变量的额外数据也许并不可行。但是，如果不存在这些限制的话，那么，毫无疑问，这一补救措施肯定是可行的。

10.8.3 重新考虑模型

有些时候，所选模型并未加以仔细考虑——或许是一些重要的变量被省略了，或许是没有正确地选择模型的函数形式。比如，在对鸡肉的需求函数中，需求函数也可能是变量之间呈线性的模型(LIV)，而不是对数线性形式的。很可能在LIV模型中共线性程度不像在对数线性型模型中那么高。

再来看对鸡肉的需求函数，我们LIV模型拟合表10-3中的数据，结果如下：

$$\hat{Y}=37.232-0.005\ 01X_2-0.611\ 2X_3+0.198\ 4X_4+0.069\ 5X_5 \tag{10-19}$$
$$t=(10.015)\quad(1.024\ 1)\quad(-3.753\ 0)\quad(3.113\ 7)\quad(1.363\ 1)$$
$$R^2=0.942\ 6\qquad\qquad \bar{R}^2=0.929\ 8$$

与方程(10-15)相比，在变量线性模型中，收入系数是统计不显著的，但猪肉价格系数却是显著的。如何解释这一变化呢？也许在收入和价格变量之间存在高度共线性。事实上，根据表10-5我们不难确认。在前面我们已经指出：在存在高度共线性的情形下，不可能准确估计各个回归系数(也就是说，标准差较小)。

在第13章中，我们将更为详细地讨论模型的设定与选择问题，需要记住的一点是：模型的不恰当设定可能是(错误的)回归模型存在共线性的原因。

10.8.4 先验信息

有些时候，某一个特定的现象，如需求函数，需要反复地调查。根据以往的研究，我们或许知道有关参数值的某些信息。而这些信息可以用于当前的样本。具体地说，假设在过去估计过的对widegets的需求函数中，收入系数为0.9，并且是统计显著的。但如前所述，利用表10-1中的数据，我们无法估计收益(作为收入的度量)对需求数量的单独影响。如果可以相信收入系数的过去值(0.9)没有多少改变的话，则我们可以重新估计方程(10-8)，结果如下：

$$\text{需求量}=B_1+B_2\text{价格}+B_3\text{收益}+u_i=B_1+B_2\text{价格}+0.9\text{收益}+u_i \tag{10-20}$$
$$\text{需求量}-0.9\text{收益}=B_1+B_2\text{价格}+u_i$$

其中，我们利用了先验信息$B_3=0.9$。

假定先验信息是“正确的”，则我们已经“解决”了共线性问题，因为在方程(10-20)的右边现在仅有一个解释变量，因此也就不会有共线性问题。为了估计方程(10-20)，我们只需将需求量的观察值减去0.9乘以相应的收益观察值并将所得的差作为解释变量，用它对价格变量进行回归。[1]

1 注意，多重共线性经常在时间序列数据中出现，因为经济变量随经济周期而变化。在这里，可用截面数据对时间序列模型中的一个或者多个参数进行估计。

从直观上看，这是一个好的方法，但是其缺陷在于要获得外生的 (extraneous)或先验的信息并不总是可行的。更为致命的是，即使我们能够获得这一信息，但要假设先验信息在当前研究的样本中依然有效，这一假设未免要求“太高”。当然，如果预期各个样本之间的收入效应不会有太大的变化，并且我们确实有关于收入系数的先验信息，那么这一补救措施是行之有效的。

10.8.5 变量变换

偶尔地，通过对模型中变量的变换能够降低共线性程度。举例来说，在对美国总消费支出的研究中，总消费支出作为总收入和总财富的函数，我们可以采取人均的形式，也就是说，作人均消费支出对人均收入和人均财富的函数。有可能在总消费函数中存在严重的共线性问题，而在人均消费函数中其共线性问题却并没有那么严重。当然，谁也不能保证这样的一个变换总能够有助于问题的解决。

为了说明简单变量变换是如何有助于减少共线性程度的，考虑如下回归结果：(根据美国1965~1980年期间的数据)：[1]

$$\hat{Y}_t = -108.20 + 0.045X_{2t} + 0.931\ X_{3t} \qquad (10\text{-}21)$$
$$t = \text{N.A.*} \quad (1.232) \quad (1.844)$$
$$R^2 = 0.989\ 4$$

*N.A.=不是有效的

其中 Y —— 进口(十亿美元)；

X_2 —— 国民生产总值(十亿美元)；

X_3 —— 消费者价格指数(CPI)。

从理论上说，进口与GNP(作为收入的一个测度)和国内价格正相关。

回归结果表明：在5%的显著水平下(双边检验)，收入和价格系数各自均不是统计显著的。[2] 但是根据F检验，我们能够拒绝零假设：两个(部分)斜率系数联合为零(验证一下)，显然表明，回归方程(10-21)存在着共线性。为了解决这一问题，萨尔瓦多(Salvatore)得到以下回归方程：

$$\frac{\hat{Y}_t}{X_{3t}} = -1.39 + 0.202\frac{X_{2t}}{X_{3t}} \qquad (10\text{-}22)$$
$$t = \text{N.A.*}$$
$$R^2 = 0.914\ 2$$

*N.A.=不是有效的

这表明，实际进口与实际收入显著正相关。这样，通过将名义变量转换为“实际”变量(也即变换原始变量)显然削弱了共线性问题。[3]

10.8.6 其他补救措施

以下补救措施只是建议性的。在文献中还有其他一些“补救措施”，例如时间序列数据和截面数据的结合(combining time series and cross-sectional data)，要素分析或者主成分分析(factor or principal componet analysis)和岭回归法(ridge regression)。但是若对这些方法详细讨

1 参见Dominick Salvatore, *Managerial Economics*，McGraw-Hill, New York,1989, pp.156-157。符号略作改动。

2 但要注意的是，根据单边t检验，在5%的显著水平下，价格系数是显著的。

3 一些作者反对按这种方式变换变量。详细情况参见 Kuh，J.R.Meyer，“Correlation and Regression Estimates When the Data Are Ratios，”*Econometrica*, pp.400-416，October 1955。另见G.S.Maddala, *Introduction to Econometrics*, Macmillan, New York, pp.172-174。

论不但会使我们偏离本书 的主题，而且也需要一些超出本书要求的统计学知识。

10.9 小结

古典线性回归模型的一个重要假定是：解释变量之间不存在完全的线性关系，或称多重共线性。尽管在实践中很少有完全多重共线性，但接近或者高度多重共线性的情形却经常发生。因此，在实践中多重共线性这一术语是指两个或者多个变量高度线性相关。

多重共线性引起的后果有：在存在完全多重共线性时，我们无法估计单个回归系数或者它们的标准差。虽然在高度多重共线性时，可以估计单个回归系数，并且普通最小二乘估计量仍保持最优线性无偏估计的性质。但是与其系数值相比，一个或者多个系数的标准差会很大，从而导致 t 值变小。因此，根据估计的 t 值，我们认为系数并不显著不为零。换句话说，我们无法估计那些 t 值较低的变量的边际或者单个贡献：在多元回归模型中，解释变量的斜率是偏回归系数，它测度了在其他变量不变的情况下，该变量对应变量的 (边际或者单个)影响。然而，如果只是为了比较精确地估计一组系数，则只要共线性不是完全的，就总能够办到。

本章我们介绍了几种检测多重共线性的方法，并指出它们各自的优缺点。我们也讨论了各种用来“ 解决 ”多重共线性问题的补救措施并指出它们各自的优劣之处。

由于多重共线性是样本的特征，因此我们无法预知哪种检测多重共线性的方法或者哪种补救措施在任何情况下均适用。

习题

10.1 什么是共线性?什么是多重共线性?

10.2 完全和不完全多重共线性的区别是什么?

10.3 在某物体重量对高度的回归模型中 (一个高度用英尺度量，另一个高度用英寸度量)，直观地解释为什么普通最小二乘法无法估计该回归方程的系数。

10.4 考虑模型

$$Y_i=B_1+B_2X_i+B_3X_i^2+B_4X_i^3+u_i$$

其中，Y=生产的总成本，X=产出。“ 既然 X^2 和 X^3 是 X 的函数，则该模型中存在着共线性。”你认为对吗？为什么？

10.5 参考方程(7-21)，(7-22)，(7-25)和(7-27)。设 $x_{3i}=2x_{2i}$，说明为什么无法估计这些方程。

10.6 不完全多重共线性的理论后果是什么？

10.7 不完全多重共线性的实际后果是什么？

10.8 什么是方差膨胀因素(VIF)？根据式(10-14)，你能说出VIF的最小可能值和最大可能值是多少吗？

10.9 补全下列各句。

(a) 在存在接近多重共线性的情况下，回归系数的标准差会趋于_____而 t 值会趋于_____。

(b) 在存在完全多重共线性的情况下，普通最小二乘估计量是_____，其方差_____。

(c) 在其它情况不变条件下，VIF越高，则普通最小二乘估计量的_____越高。

10.10 判断正误并说明理由。

(a) 尽管存在着完全多重共线性，普通最小二乘估计量仍然是最优线性无偏估计量

(BLUE)。 (b) 在存在高度多重共线性的情况下，无法估计一个或多个偏回归系数的显著性。(c) 如果辅助回归表明某一 R_i^2 较高，则表明一定存在高度共线性。 (d) 较高的两两相关系数并不一定表明存在着高度多重共线性。 (e) 如果分析的目的仅仅是为了预测，则多重共线性并无妨碍。

10.11 在用诸如失业、货币供给、利率、消费支出等经济时间序列数据进行回归分析时，常常怀疑存在多重共线性，为什么？

10.12 考虑下面的模型：

$$Y_i=B_1+B_2X_t+B_3X_{t-1}+B_4X_{t-2}+B_5X_{t-3}+u_t$$

式中 Y —— 消费；

X —— 收入；

t —— 时间。

该模型表明：t 时期的消费支出是该期收入以及前两期收入的线性函数。这类模型称为分布滞后模型(distributed lag models)，也称为动态模型(dynamic models)(也就是说，模型涉及时间的变化)。(我们将在第14章中简要讨论这类模型)。

(a) 你是否预期这类模型中存在多重共线性，为什么？

(b) 如果你怀疑存在多重共线性，那么如何“消除”它呢？

10.13 考虑如下数据集：

Y	-10	-8	-6	-4	-2	0	2	4	6	8	10
X_2	1	2	3	4	5	6	7	8	9	10	11
X_3	1	3	5	7	9	11	13	15	17	19	21

假设你想作 Y 对 X_2 和 X_3 的多元回归，

(a) 你能估计模型的参数吗？为什么？ (b) 如果不能，你能够估计哪个参数或者参数的组合？

10.14 表10-6给出了美国1971~1986年期间的年数据。

表 10-6

年度	Y①	$X_2$②	$X_3$③	$X_4$④	$X_5$⑤	$X_6$⑥
1971	102 27	112.0	121.3	776.8	4.89	79 367
1972	108 72	111.0	125.3	839.6	4.55	82 153
1973	113 50	111.1	133.1	949.8	7.38	85 064
1974	877 5	117.5	147.7	103 8.4	8.61	86 794
1975	853 9	127.6	161.2	114 2.8	6.16	85 846
1976	999 4	135.7	170.5	125 2.6	5.22	88 752
1977	110 46	142.9	181.5	137 9.3	5.50	92 017
1978	111 64	153.8	195.3	155 1.2	7.78	96 048
1979	105 59	166.0	217.7	172 9.3	10.25	98 824
1980	897 9	179.3	247.0	191 8.0	11.28	99 303
1981	853 5	190.2	272.3	212 7.6	13.73	10 0397
1982	798 0	197.6	286.6	226 1.4	11.20	99 526
1983	917 9	202.6	297.4	242 8.1	8.69	100 834
1984	103 94	208.5	307.6	267 0.6	9.65	105 005
1985	110 39	215.2	318.5	284 1.1	7.75	107 150
1986	114 50	224.4	323.4	302 2.1	6.31	109 597

① Y —— 售出新客车的数量/千辆，未作季度调整；

② X_2 —— 新车，消费者价格指数，1967=100，未作季度调整；

③ X_3 —— 所有物品所有居民的消费者价格指数，1967=100，未作季度调整；

④ X_4 —— 个人可支配收入(PDI)/10亿美元，未作季度调整；
⑤ X_5 —— 利率，%，金融公司直接支付的票据利率；
⑥ X_6 —— 城市就业劳动力/千人，未作季度调整；
数据来源：《商业统计》，1986年，《当代商业概览》增刊，美国商业部。

考虑下面的客车总需求函数：

$$\ln Y_t=B_1+B_2\ln X_{2t}+B_3\ln X_{3t}+B_4\ln X_{4t}+B_5\ln X_{5t}+B_6\ln X_{6t}+u_t$$

其中，ln表示自然对数

(a) 同时引入价格指数X_2和X_3的原因是什么？

(b) 为什么在需求函数中引入“城市就业劳动力”？

(c) 本模型中利率的作用是什么？

(d) 你如何解释各部分斜率系数的经济意义？

(e) 求上述模型的普通最小二估计值。

10.15 习题10.14中是否存在多重共线性？你是如何知道的？

10.16 如果习题10.14中存在共线性问题，估计各辅助回归方程，并找出哪些变量是高度共线性的。

10.17 继续前一题，如果存在严重的共线性，你会除去哪一变量，为什么？如果除去一个或多个变量，你可能会犯哪类错误？

10.18 在除去一个或多个解释变量后，最终的客车需求函数是什么？这个模型在哪些方面好于包括所有解释变量的原始模型。

10.19 你认为还有其他哪些变量可以更好地解释美国的汽车需求？

10.20 R. Leighton Thomas在研究英国1961～1981年期间砖、瓷、玻璃和水泥工业的生产函数时，得到如下结果：[1]

(1) $\log \hat{Q}=-5.04+0.887\log K+0.893\log H$

se=(1.40) (0.087) (0.137) $R^2=0.878$

(2) $\log Q=-8.57+0.027\ 2t+0.460\log K+1.285\log H$

se=(2.99) (0.020 4) (0.333) (0.324) $R^2=0.889$

式中 Q —— 固定成本的生产指数；
K —— 1975年重置成本的总资本存量；
H —— 工作的小时数；
t —— 时间趋势，作为技术的一种测度方法。
括号中的数字是估计的标准差。

(a) 解释这两个回归方程。 (b) 在回归方程1中验证在5%的显著水平下，部分斜率系数是统计显著的。 (c) 在回归方程2中验证在5%的显著水平下，t和$\log K$的系数各自均是统计不显著的。 (d) 如何解释模型2中变量$\log K$的不显著性？ (e) 如果得知t和K之间的相关系数为0.980，则你能够得出什么结论？ (f) 在模型2中即使t和K各自都是不显著的，你是接受还是拒绝假设：模型2中所有的部分斜率系数同时为零？你使用何种检验？ (g) 在模型1中，规模收益是什么？

10.21 确定方程(10-12)和(10-13)。(提示：找出X_2和X_3之间的相关系数，例如R^2。)

---

1 R. Leighton Thomas，*Introductory Econometrics : Theory and Applications*，Longman ,London,1985，pp. 244-246。

10.22 表10-7给出了以美元计算的每周消费支出(Y)，每周收入(X_2)和财富(X_3)等的假想数据。

表10-7 每周消费支出(Y)，每周收入(X_2)和财富(X_3)假想数据

Y	X_2	X_3
70	80	810
65	100	100 9
90	120	127 3
95	140	142 5
110	160	163 3
115	180	187 6
120	200	225 2
140	220	220 1
155	240	243 5
150	260	268 6

(a) 作Y对X_2和X_3的普通最小二乘回归。 (b) 这一回归方程中是否存在着共线性？你是如何知道的？ (c) 分别作Y对X_2和X_3的回归，这些回归结果表明了什么？ (d) 作X_3对X_2的回归。这一回归结果表明了什么？ (e) 如果存在严重的共线性，你是否会除去一个解释变量？为什么？

10.23 利用表10-1所给数据估计方程(10-20)，并比较你的结果。

10.24 验证表10-5中所有R^2值是统计显著的。

第11章

异 方 差

古典线性回归模型(CLRM)的一个重要假设是进入总体回归方程(PRF)的随机扰动项u_i同方差，也就是说，它们具有相同的方差σ^2。如果不是这样——u_i的方差为σ_i^2，也即方差随观察值不同而异(注意σ^2的下标)——这就是异方差性，或称非同方差，非常量方差。

古典线性模型强调同方差假定，但在实际操作中我们无法保证这一假设总能够满足。因此，本章主要讨论当同方差假定不满足时会发生什么情况。特别地，我们探求下列问题的答案：

(1) 异方差的性质是什么？

(2) 异方差的后果是什么？

(3) 如何检验异方差的存在？

(4) 如果存在异方差，有哪些补救措施？

11.1 异方差的性质

为了更好地解释同方差和异方差的区别，我们来看一个双变量线性回归模型，其中，应变量Y是个人储蓄，解释变量X是个人可支配收入或税后收入(PDI)。先来看图11-1。

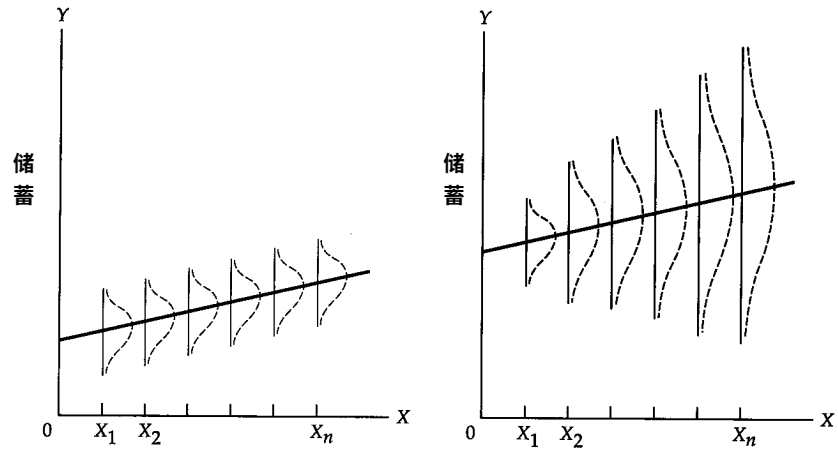


图 11-1

图11-1(a)表明，随着个人可支配收入的增加，储蓄的平均水平也增加，但是储蓄的方差在所有可支配收入水平上保持不变。(回忆一下，总体回归方程给出了对应解释变量某一给定水平下的应变量的均值。)这是同方差(homosce dasticity)或者等方差(equal variance)情形。另一方面，如图11-1(b)所示，尽管随着个人可支配收入的增加，平均储蓄水平也增加，但在各个 PDI水平上，储蓄的方差并非保持 —— 随着个人可支配收入的上升而增加。这就是异方差(heteroscedasticity)或者非同方差(unequal variance)情形。换句话说，图11-1(b)表明，平均而言，高收入者比低收入者储蓄得多，但高收入者的储蓄变动也较大。这种情况在现实中不仅是可能的，而且我们只要稍稍看一下美国储蓄与收入统计数据，就很容易证实这一点。毕竟，对低收入者而言，他们能够剩下用作储蓄的收入是非常有限的。因此，在收入对储蓄的回归分析中，预期与高收入家庭有关的误差(也就是u_i的方差)的方差比与低收入家庭有关的误差的方差要大一些。

用符号表示异方差为：

$$E(u_i^2) = \sigma_i^2 \tag{11-1}$$

再次到提醒注意σ^2的下标，表明u_i的方差不再是固定的，而是随着观察值的不同而变化。

研究人员发现，异方差问题多存在于横截面数据(cross-sectional data)中而非时间序列数据。[1]在横截面数据中，我们通常处理的是一定时间点上总体单位，例如个别消费者或是其家庭成员、公司、行业，或者区域上分区、县、州或城市等等。而且，这些单位具有不同的规模，如小公司、中等公司或者大公司，或是低收入、中等收入或高收入。换言之，可能存在着规模效应(scale effect)。而在时间序列数据中，变量趋于具有相似的数量等级，因为我们通常收集某一时期同一个实体的数据。例如，从 1960年到1990年的国民总产值、储蓄、失业率等等。

我们用两个具体的例子来说明异方差问题。

例11.1 放松管制后纽约股票交易所(NYSE)的经纪人佣金。

1975年四五月间，债券交易委员会 (Securities and Exchange Commission，SEC) 废除了对于纽约股票交易所股票交易固定佣金率的规定，允许股票经纪人在竞争的基础上索取佣金。表 11-1给出了从1975年4月到1978年12月间经纪人对机构投资者索要的平均每股佣金的季度数据。

表11-1 纽约股票交易所佣金率趋势

(单位：美分)

月份	年度	X_1	X_2	X_3	X_4
4月	1975	59.6	45.7	27.6	15.0
6月		54.5	36.8	21.3	12.1
9月		51.7	34.5	20.4	11.5
12月		48.9	31.9	18.9	10.4
3月	1976	50.3	33.8	19.0	10.8
6月		50.0	33.4	19.5	10.9
9月		46.7	31.1	18.4	10.2
12月		47.0	31.2	17.6	10.0
3月	1977	44.3	28.8	16.0	9.8
6月		43.7	28.1	15.5	9.7
9月		40.4	26.1	14.5	9.1

1 严格地说，这是不正确的。在所谓的 ARCH(自回归条件异方差)模型中，在时间序列数据中也能观察到异方差。这是一个较深的专题，在本书中不予讨论。有关 ARCH模型的讨论，可参见 G.S. Maddala，*Introduction to Econometrics*，Macmillan，New York，1988，pp.218-219。

(续)

(续)

月份	年度	X_1	X_2	X_3	X_4
12月		40.4	25.4	14.0	8.9
3月	1978	40.2	25.0	13.9	8.1
6月		43.1	27.0	14.4	8.5
9月		42.5	26.9	14.4	8.7
12月		40.7	24.5	13.7	7.8

名称	n	平均值	标准差	方差	最小值	最大值
X_1	16	46.5	5.676 7	32.223	40.2	59.6
X_2	16	30.673	5.501 6	30.268	24.5	45.7
X_3	16	17.444	3.723 4	13.864	13.7	27.6
X_4	16	10.094	1.783 4	3.1806	7.8	15.0

注： X_1——佣金率，美分/股(0至199股)

X_2——佣金率，美分/股(200至299股)

X_3——佣金率，美分/股(1000至9999股)

X_4——佣金率，美分/股(10000股以上)

数据来源： S. Tinic and R. West，“The Securities Industry Under Negotiated Brokerage Commissions in the Structure and Performance of NYSE Member Firms”，*The Bell Journal of Economics*，vol.11，no.1，Spring 1980.

注意表中两个有趣的特征。放松管制以来，佣金率有下降的趋势。然而，更令人感兴趣的是平均佣金率存在着显著的不同，并且对四类机构投资者所索取的佣金的方差也存在着显著的差异。最小的机构投资者 (那些股票交易量介于0至199股的机构投资者)平均每股需付46.5美分，其方差为32.22，而最大的机构投资者平均每股只需付10.1美分，方差只有3.18。读者能够从图11-2中清楚地看到。

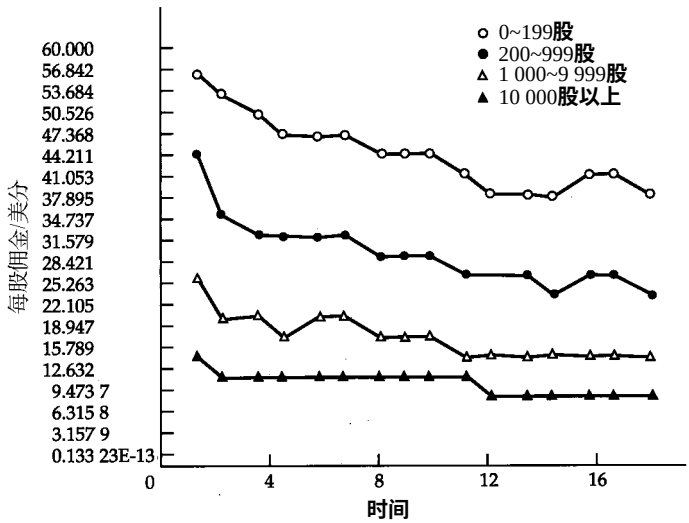


图11-2 纽约股票交易所每股佣金

注： 根据表11-1提供的数据

(续)

如何来解释这种差异呢？显然，这里存在着规模效应——交易量越大，则交易的总成本就越低，因而平均成本也就越低。经济学家们会说，表 11-1中所给的经纪行业数据中存在着规模经济(但这也并不一定，参见例11.7和本章第6节)。但即使在经纪行业中存在着规模经济，那为什么四类的佣金率的方差会不同呢？换句话说，为什么存在着异方差？为了吸引大机构投资业务，比如养老基金、共同基金等，经纪公司相互之间竞争激烈，因而他们索要的佣金率并没有多大差异。小的机构投资者也许就没有大机构投资者那样的谈判能力，因而他们支付的佣金率也就存在着较大的差异。这或许能够解释在表11-1数据中所观察到的异方差(当然，也许还有其他的原因)。

现在如果我们想建立一个回归模型来解释佣金率对股票交易数量(和其他变量)的函数，那么与高交易量客户相关的误差项方差将会低于与低交易量客户相关的误差项方差。

例11.2 美国工业的研究与发展费用支出、销售和利润。

我们给出一个纯横截面数据的例子(可能存在异方差)，考虑表11-2所给出的数据。[1] 包括美国18个行业1988年的销售、利润及研究与发展(R&D)支出数据，所有数据单位都是百万美元。由于每个行业包括若干不同子类，各子类所包括公司的规模又各不相同，如果作研究和发展支出对销售量和利润的回归，则很难保证同方差假定，其原因在于行业分类的多样性。

现在假设我们想知道研究与发展支出与销售的相关关系。我们考虑如下模型：

$$\text{R\&D}_i = B_1 + B_2 \text{销售额}_i + u_i \tag{11-2}$$

先验地，我们预期两个变量呈正相关关系(为什么？)，这与图11-3中的散点图看似一致。该图描绘了R&D支出与销售的关系。

对方程(11-2)利用普通最小二乘法，得到回归结果如下：

$$\widehat{\text{R\&D}}_i = 192.99 + 0.031\,9 \text{销售额}_i$$
$$\text{se} = (990.99)\ (0.0083) \tag{11-3}$$
$$t = (0.194\,8)\ (3.843\,4) \quad r^2 = 0.478\,3$$

图11-3描绘了这条回归线。

表11-2 1988年美国研究与发展(R&D)支出费用

(单位:百万美元)

序号	行 业	销售额	R&D费用支出	利 润
1	容器与包装	6 375.3	62.5	185.1
2	非银行金融机构	11 626.4	92.9	1 569.5
3	服务行业	14 655.1	178.3	274.8
4	金属与采掘业	21 896.2	258.4	2 828.1
5	住房与建筑业	26 408.3	494.7	225.9
6	一般制造业	32 405.6	1 083.0	3 751.9
7	闲暇时间行业	35 107.7	1 620.6	2 884.1
8	纸与林产品行业	40 295.4	421.7	4 645.7

1 与该表所给数据相比，表11-1的数据既有时间序列数据又有横截面数据：四个交易类的任一给定月度数据都是横截面数据，而任给一类从1975年4月到1978年12月的数据又都是时间序列数据。

(续)

(续)

序号	行 业	销售额	R&D费用支出	利 润
9	食品行业	70 761.6	509.2	5 036.4
10	健康护理业	80 552.8	6 620.1	13 869.9
11	宇航业	95 294.0	3 918.6	4 487.8
12	消费品	101 314.1	1 595.3	10 278.9
13	电器与电子产品	116 141.3	6 107.5	8 787.3
14	化学工业	122 315.7	4 454.1	16 438.8
15	聚合物	141 649.9	3 163.8	9 761.4
16	办公设备与计算机	175 025.8	13 210.7	19 774.5
17	燃料	230 614.5	1 703.8	22 626.6
18	汽车行业	293 543.0	9 528.2	18 415.4

注：行业是按销售额递增的次序排列的。

资料来源：*Business Week* ,Special 1989 Bonus Issue ,R&D Scorecard，pp.180-224.

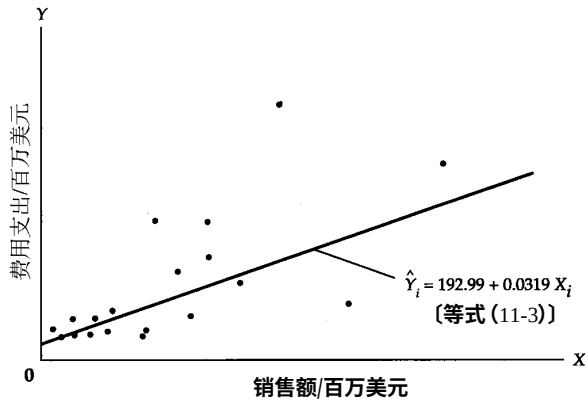


图11-3 1988年美国工业研究与发展(R&D)费用支出与销售额

从图11-3可以看出：平均而言，R&D支出费用随着销售额的增加而增加。但是，引人注意的是：随着销售的增加，R&D支出费用的变动幅度也增大了，也即存在着异方差。如果我们把从回归方程中得到的残差对各个观察值作图 (图11-4)，则可以更明显地看到这一点：因为观察值是按照销售额升序排列的，所以这就等于间接地将残差对销售额作图。

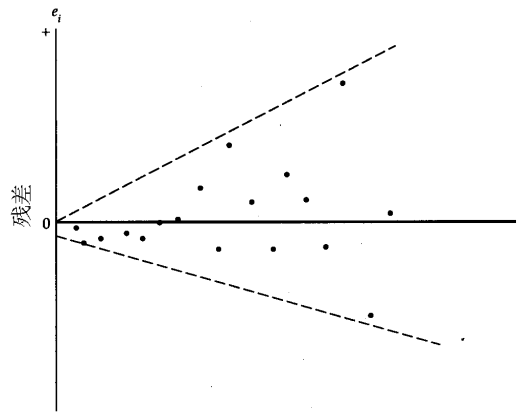


图11-4 R&D回归方程(11-3)中的残差

从图11-4中可以很清楚地看到，残差的绝对值随着销售额的增加而增加。这再一次表明，在这个例子中同方差假定是无法站住脚的。因此，我们将不得不重新考虑回归方程 (11-3)，因为该方程是建立在同方差假定基础之上的。我们将在本章的后半部分解决这个问题。

特别需要指出的是，尽管残差 e_i 与扰动项 u_i 很相像，但二者并不相同。因此，根据可观察到的 e_i 的变化，我们并不能直接推出 u_i 的方差也是变化的。但是，随后我们将会看到，在实际中，我们很少能够观察到 u_i，因此，我们只能够处理 e_i。换句话说，通过对 e_i 进行图形检验，来推得有关 u_i 的图形。

假设在R&D回归方程中，根据图11-3和图11-4，我们认为出现了异方差情况。接着该怎么办呢？是不是基于同方差假定所得到的式(11-3)回归结果毫无用处呢？[1] 为了回答这一问题，我们必须知晓在异方差情况下普通最小二乘法会出现什么异常。下面我们讨论这个问题。

11.2 异方差的后果

我们知道在古典线性回归模型的假设下，普通最小二乘法估计量是最优线性无偏估计量，也就是说，在众多线性无偏估计量中，最小二乘估计量具有最小方差性 —— 它是有效估计量。现在假定我们解除同方差假定，允许扰动方差随观察值而异，但其他假定保持不变。下面给出异方差的后果，证明从略。[2]

(1) OLS估计量仍然是线性的。

(2) OLS也是无偏的。

(3) 但它们不再具有最小方差性。也就是说，它们不再是有效的。即使对大样本也如此。简言之，无论是小样本还是大样本，OLS估计量都不再是最优线性无偏估计量。

(4) 根据常用估计OLS估计量方差的公式得到的方差通常是有偏的。先验地，我们无法辨别是正的偏差还是负的偏差。如果OLS高估了估计量的真实方差，则会产生正的偏差，而如果OLS低估了估计量的真实方差，则会产生负的偏差。

(5) 方差的产生是由于 $\hat{\sigma}^2$，也即 $\sum e_i^2/\text{d.f.}$ 不再是真实 σ^2 的无偏估计量。[注：d.f.(自由度)在双变量模型中是 $(n-2)$，在三变量模型中是 $(n-3)$，等等]别忘了，在计算OLS估计量的方差时已经用了 $\hat{\sigma}^2$。

(6) 因此，建立在 t 分布和 F 分布之上的置信区间和假设检验是不可靠的。如果仍用传统的假设检验方法，则有可能得出错误的结论。

简言之，在存在异方差的情况下，通常所用的假设检验已经不再可靠，因为有可能得出错误的结论。

回到模型(11-3)所给出的R&D回归结果，如果存在着异方差或是有迹象表明存在异方差 (关于异方差存在性的正规检验方法将在本章第 3节予以讨论)，那么，在解释回归结果时必须非常小心。从(11-3)回归可以看出：“显然”，销售变量的系数显著不为零，因为其 t 值为3.83，在1%的显著水平下，是“显著地”的 (对于自由度为16，在0.01显著水平下，单边 t 临界值为2.921)。实际上，ρ值小于0.000 1。但如果确实存在异方差，那么我们就不能相信估计得到的标准差，0.008 3，因而也就不能相信计算的 t 值。这就促使我们去检查是否确实存在着异方差问题。

---

1 作为实际问题，在进行回归时，我们通常假定古典线性回归模型所有的假设都得到满足。只是在我们检查回归结果时，我们才开始寻找一些“线索”，这些线索或许能够告诉我们古典线性回归模型的一个或者多个假定也许是不满足的。这并不完全是一个坏的策略，何必“对礼物吹毛求疵”呢？

2 证明可参阅：Damond.N.Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，Chap.11。

上述讨论表明：异方差是一个潜在的严重问题，因为它可能破坏常用的 OLS估计以及假设检验过程。因此，在具体的研究中，尤其是涉及横截面数据时，很重要的一点是首先要判断是否存在异方差问题。

但是，在讨论异方差的检验之前，我们应该知道为什么在异方差的情形下 OLS估计量是无效的。

考虑双变量回归模型。回顾第 5章的内容，在运用普通最小二乘法的过程中我们要使残差平方和(RSS)最小：

$$\sum e_i^2 = \sum (Y_i - b_i - b_2 X_i)$$

现在考虑图11-5

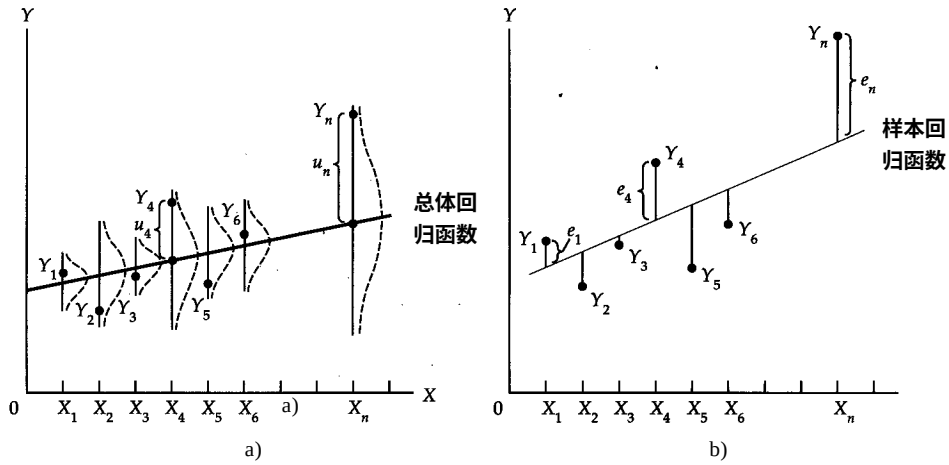


图 11-5

该图形描绘了某一假设总体 Y对变量X的所选值间的关系。从图中可以看出，给定 X，对应每一(子)总体Y的方差是不同的，这表明存在异方差。假设对应每一个 X值随机地选取一个Y值。将所选的Y值圈起来。根据方程(5-13)可知，利用普通最小二乘法，每一个 e_i^2都有同样的权重，无论它是来自于一个较大方差的总体还是来自于一个较小方差的总体 (比较点Y_n和点Y_1)。这么做似乎并不明智，我们应该给那些取自较小方差总体的观察值以更大的权重，而给那些取自较大方差总体的观察值以更小的权重。这能够使我们更为精确地估计总体回归函数。这就是加权最小二乘法(weighted least squares)，稍后我们讨论。

11.3 异方差的检验：如何知道存在异方差问题

尽管理论上容易列举异方差的后果，但对具体情况检验异方差问题却绝非易事。这一点容易理解，因为只有当我们有了与所选 X相对应的整个 Y总体时，才能知道 σ_i^2(如表5-1给出的widgets一例的假设的总体)。然而，不幸的是，我们很少能够得知整个总体。更一般地，我们仅仅知道一个样本。典型地，我们仅有与给定变量 X值相对应的单独的一个 Y值。而根据单独的这个Y值无法确定对应于给定X值的Y的条件分布的方差。[1]

为了更清楚地看到这一点，我们再来看表 11-2给出的R&D数据。对每一个行业，只有一个R&D数据。仅从这个R&D数据，根本无法求出该行业的 R&D的方差。例如，对于电器和电子

1 注意，给定X值，u的方差和Y的方差是一样的。换句话说，u的条件方差(对应于给定的X值)与Y的条件方差是相同的，参见第6章脚注1。

行业，其总的R&D支出约为6 107(百万美元)。但是，该行业中包括了171个公司。除非我们知道关于每个公司的信息，否则我们无法说出该行业 R&D支出费用的变化。也就是说，如果将6 107除以171得到大约38(百万美元)作为该行业的平均R&D支出费用，那么我们无法确定R&D支出费用是如何围绕这一均值分布的。换句话说，我们无法确定该行业的 R&D支出费用的方差是多少。对于表11-2中的其他17个行业也是如此。

因此，根据表11-2的数据，如果作R&D对销售的回归，正如方程(11-3)那样，我们无法知道与每个观察值(也就是每个R&D数据)有关的σ_i^2值：回归方程(11-3)是基于假设——表11-2中每个观察值都具有同样的(也就是同方差)方差σ^2。而这一方差是通过$\hat{\sigma}^2 = \sum e_i^2/(n-2)$估计得到的。这里，我们的任务就是判定这一假设是否正确。

现在我们处于“进退维谷”的地步。如果存在着异方差但我们却假定了不存在，那么根据普通最小二乘法可能得出误导性结论，因为OLS估计量不是最优线性无偏估计量。但是由于大多数数据来自同一个样本，因而我们又没有其他的方法能够求出与每个观察值有关的真实误差σ_i^2。如果我们能够找出真实的σ_i^2，就可能解决异方差问题(参见本章第4节)。那么，经济计量学家能做什么呢？

与多重共线性的情形相同，并没有严格的规则用来检验异方差；我们所有的只是一些检测工具，这些检测工具可以帮助我们检验异方差。下面介绍其中的一些检测工具。

11.3.1 根据问题的性质

所考察问题的性质往往提供是否存在异方差的信息。例如，根据普雷斯 (Paris)和霍撒克(Houthakker)[1]关于家庭预算的开创性著作，他们发现在消费对收入的回归中，残差方差随收入的增加而增加。现在一般相类似的调查研究通常假设不同扰动项的方差是不等的。事实上，在涉及不均匀单位的横截面数据中，异方差可能是常有的情况而不是例外。例如，在与销售、利率等相关的投资支出的横截面数据分析中，如果把小、中和大型的公司聚集在一起加以抽样，就很可能存在异方差。类似地，在对与产出相关的平均成本的横截面数据研究中，也很可能存在异方差，如果样本包括了小、中和大型公司。(参见例11.8)

11.3.2 残差的图形检验

在回归分析中，常常对拟合回归方程中的残差进行分析。将残差对其相应的观察值描图，或是对一个或多个解释变量描图，或是对估计的 Y_i的均值，$\hat{Y}_i$ 描图。根据这样的残差图(residual plot)可以为我们判断古典线性回归模型中的一个或者多个假定是否满足提供线索，我们将简要地加以说明。

我们已看到了图11-4，该图描绘了从回归方程(11-3)中所得的残差与销售的关系。从图中可以看出，残差的(绝对)值随销售量的增加而增加。实际上，该图是一个放大器的形式。或许数据中存在着异方差。

有时我们不是将残差对销售描图，而是将残差的平方e_i^2对销售描图。尽管e_i^2与u_i^2不同，它经常可以用它来替代后者，尤其对大样本。[2]

在利用R&D回归方程(11-3)的数据将残差平方对销售描图之前，我们先考虑一下当将e_i^2对变量X描图时可能遇到的异方差的各种模式。参见图 11-6。在图11-6a中，变量X与e_i^2之间没有可观察到的系统模式，表明数据中可能不存在异方差。另一方面，从图 11-6b到e可以看出残差

1 S .J.Paris，H.S.Houthakker，The Analysis of Family Budgets，Cambridge University，New York，1955.

2 e_i^2与u_i^2之间的关系，参见 E. Malinvaud，*Statistical Methods of Econometrics*，North-Holland, Amsterdam,1970,pp.88-89.

平方与解释变量 X 之间的系统关系；例如，图 11-6c表明，两者之间存在着线性关系。而图 11-6d和e则表明存在着四次方关系。因此，在实践中，如果残差平方呈现出图 11-6b到e中的任意一种，则数据中很可能存在着异方差。

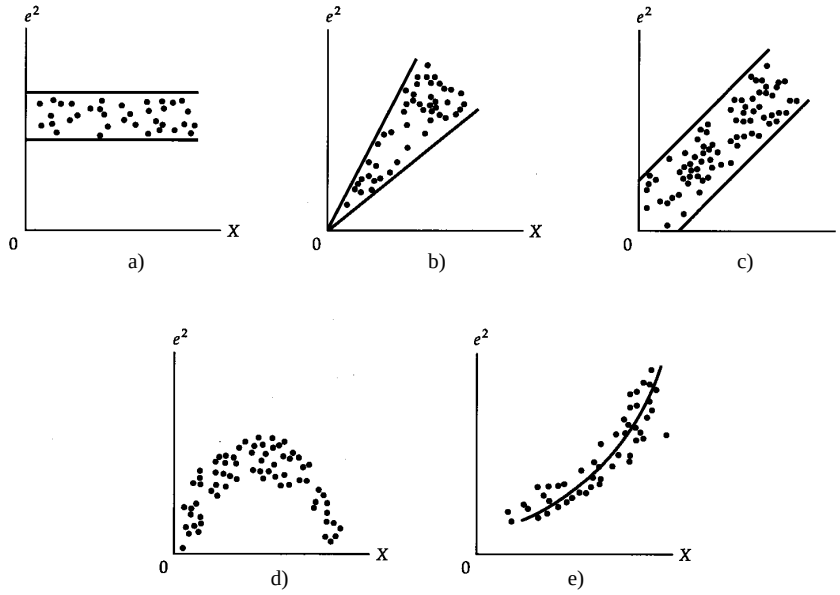


图11-6 e^2模式

注意，上述散点图只不过是一个检测工具，一旦怀疑存在异方差，再继续分析时我们应该更为谨慎。

同时，提出一些实际的问题。比方说，假设现有一个包括四个变量的多元回归方程。那么我们该如何继续下去呢？最直观的方法是将 e_i^2 对每个变量描图。可能只有一个变量表现出图 11-6中(b到e)的某个模式。有时我们可以寻找一个捷径。将 e_i^2 对 Y 的估计均值 $\hat{Y}_i$ 描图，而不是将 e_i^2 与每个变量描图。由于 $\hat{Y}_i$ 是 X_s 的线性组合(为什么？)，e_i^2 对 $\hat{Y}_i$ 的散点图可能会呈现出图 11-6(b到e)某种模式，表明数据中可能存在异方差。这就避免了将残差平方对单个变量描图的烦琐过程，尤其是当模型中的解释变量很多时。

假定我们将 e_i^2 对一个或者多个解释变量或是 $\hat{Y}_i$ 描图，进一步假定作出的散点图表明存在着异方差。那么接下来该怎么办呢？在 11.4节中，我们将介绍如果 e_i^2 与解释变量或是 $\hat{Y}_i$ 相关，如何从原始数据进行变换，使得变换后的数据不存在异方差。

我们回到R&D一例。在图11-7中，将从回归方程(11-3)中估计得到的残差平方对模型中的解释变量，销售描图。[1]

该图与图11-6 (b)很相似，这清楚地表明了残差平方与销售是系统相关的。该散点图表明在R&D回归方程(11-3)中可能存在异方差问题。

11.3.3 帕克检验(Park test)[2]

上面给出的图形检验比较直观，我们可以加以规范。如果存在着异方差，异方差方差 σ_i^2 可

1 注意，我们是将 e_i^2 而不是 e_i 对 X 或者 $\hat{Y}_i$ 描图，正如我们在第6章中脚注13所指出的那样，e_i 与 X 以及 $\hat{Y}_i$ 之间都是零相关。

2 R.E.Park，“Estimation with Heterosceastic Errors Terms”，*Econometrica* , vol.34，no.4，October 1966，p.888.

能与一个或者多个解释变量系统相关。为了弄清楚情况是否果真如此，可以作 σ_i^2 对一个或者多个解释变量的回归。例如在双变量模型中，我们可以运行下面的回归方程：

$$\ln \sigma_i^2 = B_1 + B_2 \ln X_i + v_i \tag{11-4}$$

其中，v_i 是残差项。这就是帕克检验。这里所选择的特殊函数形式 (11-4)是为了方便起见。

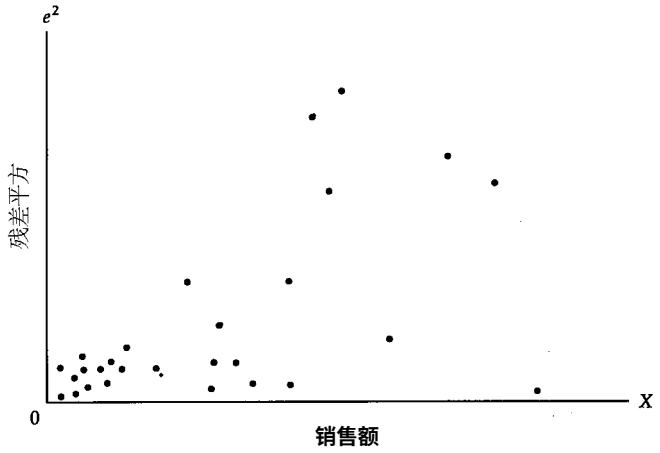


图11-7 e_i^2 与销售，R&D回归方程(11-3)

不幸的是，回归方程(11-4)是不可操作的，因为我们并不知道异方差方差 σ_i^2。如果知道的话，可以很容易解决异方差的问题(参见11.4节)。帕克建议用 e_i 来代替 u_i，运行如下回归方程：

$$\ln e_i^2 = B_1 + B_2 \ln X_i + v_i \tag{11-5}$$

如何来获得 e_i^2 呢？我们可以从原始的回归方程中获得 e_i^2 的值，比如模型(11-3)。

帕克检验的步骤如下：

(1) 作普通最小而乘回归，不考虑异方差问题。

(2) 从原始回归方程中得残差 e_i，并求其平方，再取对数形式(计算机能够做到这一点)。

(3) 利用原始模型中的一个解释变量作形如式(11-5)的回归；如果有多个解释变量，则对每个解释变量都做形如式(11-5)的回归。或者作 e_i^2 对 Y 的估计值 $\hat{Y}_i$ 的回归。[1]

(4) 检验零假设 $B_2=0$；也即不存在异方差。如果 $\ln e_i^2$ 和 X_i 之间是统计显著的，则拒绝零假设：不存在异方差，在这种情况下需要采取一些补救措施，我们将在 11.4节中讨论。

(5) 如果接受零假设，则回归方程中的 B_1 可以解释为同方差 σ^2。

例11.3 R&D回归与帕克检验。

我们来解释R&D回归方程(11-3)。把从这个回归方程中得到的残差用于回归模型(11-5)，得到如下结果：

$$\ln \hat{e}_i^2 = 5.687\ 7 + 0.701\ 4 \ln \text{销售额}_i$$
$$\text{se} = (6.635\ 2)\ \ (0.603\ 3) \tag{11-6}$$
$$t = (0.857\ 2)\ \ (1.162\ 6)\ \ r^2 = 0.077\ 9$$

显然，在5%的显著水平下(单边检验)，估计的斜率系数是统计不显著的。

我们能否因此而接受假设：R&D数据中不存在有异方差问题。我们还不能这么

1 在运行回归方程(11-5)时，需要考虑选择合适的函数形式。有些时候，作 e_i^2 对 X_i 的回归可能是合适的，有些时候，作 $\ln e_i^2$ 对 X_i 的回归可能是合适的。

(续)

快下结论。因为帕克所选择的特殊函数形式，回归方程 (11-5)，只是建议性的，其他的函数形式也许会使我们得出不同的结论 (在第13章中，我们将更多的讨论如何选择“合适”模型)。帕克检验存在一个比较严重的问题，就是在回归方程 (11-5)中，误差项v_i本身可能存在着异方差！我们又回到问题起点。在判定 R&D回归方程中不存在异方差之前，我们需要更多的检验。

11.3.4 Glejser检验(Glejser test)[1]

Glejser检验实质上与帕克检验很相似。从原始模型中获得残差 e_i之后，Glejser建议作e_i的绝对值$|e_i|$对X的回归。Glejser建议的一些函数形式如下：

$$|e_i| = B_1 + B_2 X_i + v_i \tag{11-7}$$

$$|e_i| = B_1 + B_2 \sqrt{X_i} + v_i \tag{11-8}$$

$$|e_i| = B_1 + B_2 \left(\frac{1}{X_i}\right) + v_i \tag{11-9}$$

每种情形的零假设都是不存在异方差，也即，B_2=0。如果零假设被拒绝，则表明可能存在着异方差。

例11.4 R&D回归与Glejser检验

根据回归方程(11-3)的残差估计这些模型，得到的结果如下：

$$|e_i| = 578.57 + 0.011\,9\,\text{销售额}_i$$
$$t = (0.852\,5)\quad (2.093\,1)\qquad r^2 = 0.215\,0 \tag{11-10}$$

$$|e_i| = -507.02 + 7.972\,0\sqrt{\text{销售额}_i}$$
$$t = (-0.503\,2)\quad (2.370\,4)\qquad r^2 = 0.259\,9 \tag{11-11}$$

$$|e_i| = -227\,3.7 - 19925000\left(\frac{1}{\text{销售额}_i}\right)$$
$$t = (3.760\,1)\quad (-1.617\,5)\qquad r^2 = 0.140\,5 \tag{11-12}$$

根据Glejser检验，模型(11-10)和(11-11)表明拒绝零假设：不存在异方差，因为在5%的显著水平下(双边检验)，斜率系数都是统计显著的。而模型 (11-12)表明接受同方差的假设。

对Glejser检验有一点需要特别注意：与帕克检验一样，在 Glejser所建议的回归方程中，误差项本身可能就存在异方差和系列相关问题(关于系列相关，参见第12章)。然而，对于大样本，上述模型能够很好地检测异方差问题。因此，Glejser检验可用作大样本的检测工具。在 R&D一例中，仅仅有18个观察值，因此，在解释回归方程(11-10)、(11-11)、(11-12)时必须特别地谨慎。

11.3.5 White检验(White's General Heteroscedasticity Test)[2]

怀特(White)提出了异方差一般检验方法，这一检验方法在实际中很容易应用。假定有如

1 H.Glejser,“A New Test for Heteroscedasticity,” *Journal of the American Statistical Association* (JASA), vol. 64，1969,316-323.

2 H. White.“A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test of Heteroscedasticity”, *Econometica* vol.48，no.4，1980，pp.817-818.

下模型：

$$Y_i=B_1+B_2X_{2i}+B_3X_{3i}+u_i \tag{11-13}$$

White检验步骤如下：

(1) 首先用普通最小二乘法估计回归方程(11-13)，获得残差e_i。

(2) 然后作如下辅助回归：

$$e_i^2=A_1+A_2X_{2i}+A_3X_{3i}+A_4X_{2i}^2+A_5X_{3i}^2+A_6X_{2i}X_{3i}+v_i \tag{11-14}$$

即作残差平方e_i^2对所有原始变量、变量平方以及变量交叉乘积的回归。也可以加入原始变量的更高次幂。v_i是辅助回归方程中的残差项。

(3) 求辅助回归方程(11-14)的R^2值。在零假设：不存在异方差(也就是，方程(11-14)中的所有斜率系数都为零)下，White证明了，从方程(11-14)中获得的R^2值与样本容量($=n$)的积服从χ^2分布，自由度等于方程(11-14)中解释变量的个数(不包括截距项)。

$$n\cdot R^2\sim\chi^2 \text{ d.f.} \tag{11-15}$$

在模型(11-14)中，自由度d.f.为5。

(4) 如果从方程(11-15)中得到的χ^2值超过了所选显著水平下的χ^2临界值，或者说计算χ^2值的ρ值很低，则可以拒绝零假设：不存在异方差。而另一方面，如果计算的χ^2值的ρ值很大，则不能拒绝零假设。

例11.5

为了说明White检验，我们来看习题11.18中表11-15的数据。在回归方程(11-13)中，Y代表婴儿死亡率，X_2代表人均GNP，X_3代表受初等教育占人口的百分比(作为文化程度的指标)。先验地，我们预期人均GNP和教育指标对婴儿死亡率有着负面的影响(为什么？)。回归结果如下：

$$\begin{aligned}\hat{Y}_i&=111.78-0.004\,2X_{2i}-0.489\,8X_{3i}\\ \text{se}&=(23.35)\quad(0.001\,6)\quad(0.286\,5)\\ t&=(4.79)\quad(-2.53)\quad(-1.71)\quad R^2=0.492\\ p&=(0.000)^{①}\quad(0.022)\quad(0.106)\quad \bar{R}^2=0.433\end{aligned} \tag{11-16}$$

① 表示一个很小的值。

与预期相同，X_2和X_3的系数均为负。根据单边t检验(为什么要用单边检验？)，两个斜率系数t值的p值分别为0.011和0.053(是方程(11-16)中p值的一半)。

正如习题11.18所示，表11-5给出的是20个国家的数据，其中有5个国家是属于低收入国家，5个国家是中等偏下收入国家，5个国家是中等偏上收入国家，5个国家是高收入国家。其中，收入的划分是按照世界银行的划分方法。由于我们的样本是来自经济条件差异很大的国家，因此，先验地，我们预期存在异方差。为了检验我们的预期，我们使用White检验，利用模型(11-14)。回归的结果在习题11.19中给出，回归结果表明R^2值为0.649。然后，利用方程(11-15)，得到，

$$20(0.649)=12.98 \tag{11-17}$$

在零假设：不存在异方差下，方程(11-17)所给出的值服从自由度为5的χ^2分布。得此χ^2值的ρ值约为0.023 35，这一值是很小。因此，根据White检验，我们可以下判定存在异方差。

怀特检验的一个缺陷是它太一般化了。如果有好几个解释变量的话，则在回归方程中要包括这些变量，变量的平方(或者更高次幂)以及它们的交叉乘积项，这会迅速地降低自由度。因此，在引入太多变量时，必须谨慎一些。有时，我们可以去掉变量的交叉乘积项。

除了检验异方差以外，White还提供了一种统计方法用于修正形如回归方程 (11-13)的估计标准差。我们将在本章第5节中详细讨论。

11.3.6 异方差的其他检验方法

上面只是讨论了检验异方差的一些常用的方法。下面给出了异方差的检验的一些其他方法，这里我们不在详细讨论。

(1) 斯皮尔曼(Spearman)秩相关检验(参见题11.13)。

(2) 戈德费尔德-匡特(Goldfeld-Quandt)检验。

(3) 巴特莱特(Bartlett)方差同质性检验。

(4) 匹克(Peak)检验。

(5) 布鲁尔什-培甘(Breusch-Pagan)检验。

(6) CUSUMSQ检验。

感兴趣的读者可以查阅有关这方面的参考书。[1]

11.4 观察到异方差该怎么办：补救措施

我们在前面已经看到，异方差的存在并不破坏普通最小二乘法估计量的无偏性，但是估计量却不再是有效的，即使对大样本也是如此。缺乏有效性，就使通常假设检验程序的值不可靠。因此，如果怀疑存在异方差或者已经检测到了异方差的存在，则寻求补救的措施就很重要。

例如，在R&D一例中，根据Glejser检验，表明在回归方程 (11-3)中可能存在异方差问题。如何解决这一问题呢？是否存在这样方法，将模型 (11-3)加以“变换”，使得“变换”后的模型具有同方差性？但是采取什么样的变换形式呢？答案取决于 (1)误差的真实值 σ_i^2 是已知的(2)误差的真实值 σ_i^2 是未知的。

11.4.1 加权最小二乘法(WLS)

我们考虑双变量总体回归函数

$$Y_i=B_1+B_2X_i+u_i \tag{11-18}$$

其中Y，比方说是R&D支出，X是销售量。从现在起，假设误差 σ_i^2是已知的，也就是说，每个观察值的误差是已知的。对模型作如下“变换”：

$$\frac{Y_i}{\sigma_i}=B_1\left(\frac{1}{\sigma_i}\right)+B_2\left(\frac{X_i}{\sigma_i}\right)+\frac{u_i}{\sigma_i} \tag{11-19}$$

这里将回归等式(11-18)的两边都除以“已知”的σ_i。σ_i是方差σ_i^2的平方根。

令，

$$v_i=\frac{u_i}{\sigma_i} \tag{11-20}$$

我们将v_i称作是“变换”后的误差项。v_i满足同方差吗？如果是，则变换后的回归方程 (11-19)就不存在异方差问题了。假设古典线性回归模型中的其它假设均能满足，则方程 (11-19)中各参数的OLS估计量将是最优线性无偏估计量，我们就可以按常规的方法进行统计分析了。

证明误差项v_i同方差性并不困难。根据方程(11-20)，有：

---

1 Spearman秩相关检验，Goldfeld-Quandt检验和Breush-Pagan检验参见Damodar.N.Gujarati，*Basic Econometrics*，3ed，McGraw-Hill，New York,1995，Chap11.

$$v_i^2 = \frac{u_i^2}{\sigma_i^2} \tag{11-21}$$

因此，

$$\begin{aligned} E(v_i^2) &= E\left(\frac{u_i^2}{\sigma_i^2}\right) \\ &= \frac{1}{\sigma_i^2} E(u_i^2)，由于 \sigma_i^2 是已知的。 \\ &= \frac{1}{\sigma_i^2}(\sigma_i^2)，由方程(11-1) \\ &= 1 \end{aligned} \tag{11-22}$$

显然它是一个常量。简言之，变换后的误差项 v_i 是同方差的。因此，变换后的模型 (11-19) 不存在异方差问题，我们可以用常规的 OLS方法加以估计。

在实际估计回归方程 (11-19)中，你需要给计算机一个指令，将 Y和X的每个观察值都除以“已知”的 σ_i，然后再对这些变换后的数据进行 OLS回归(现在大多数计算机软件包都能够实现这一点)。由此获得的B_1、B_2的OLS估计量就称为加权最小二乘(weighted least squares, WLS)估计量；Y和X的每个观察值都以其(异方差)标准差 σ_i 为权数(也就是说，除以 σ_i)。这种加权的过程就称为加权最小二乘法(WLS) [1] (参见习题11.14)。

11.4.2 σ_i^2 为未知时的情况

尽管从直观上看加权二乘法很简单，但有一个重要的问题：如何知道或者如何找出真实的误差方差，σ_i^2 ？前面我们已经指出，在经济计量学的研究中，有关误差方差的信息是极少的。因此，如果想要使用 WLS方法，常借助于似乎可靠的关于σ_i^2 的假设，把原来的回归模型变换为能够满足同方差假定的模型。然后才可以对变换后的模型运用 OLS法，因为WLS只不过是对变换过的数据用OLS法。 [2]

当σ_i^2 未知时，紧接着的一个实际问题就是，对于这个未知的误差方差我们要做何假设？怎样才能应用WLS法？在此，我们考虑双变量模型所出现的几种可能情况；我们也很容易将其推广到多元回归的情形。

情形1：误差与X_i成比例：平方根变换。在用常规的OLS法估计之后，我们将回归所得的残差对解释变量X描图，如果观察到图案与图 11-8相似，则表明误差与解释变量X线性相关，或者说与X成比例。

也就是，

$$E(u_i^2) = \sigma^2 X_i \tag{11-23}$$

这表明误差方差与X成比例，或者说与X_i线性相关；常数 σ^2 (注意 σ^2 没有下标)是比例因子。

在式(11-23)假定下，将模型(11-18)作如下变换：

$$\frac{Y_i}{\sqrt{X_i}} = B_1 \frac{1}{\sqrt{X_i}} + B_2 \frac{X_i}{\sqrt{X_i}} + \frac{u_i}{\sqrt{X_i}} \tag{11-24}$$

---

1 注意回归方程(11-19)的技术要点。为了估计它，需要给计算机一个指令，让回归线经过原点，因为方程(11-19)中没有“明确”的截距——方程的第一项是 $B_1(1/\sigma_i)$。但是，$(1/\sigma_i)$的斜率系数实际上是截距系数B_1。(你看出这一点了吗？)对经过原点回归的讨论，参见习题 6.15。

2 注意,在OLS中,是最小化 $\sum e_i^2 = \sum (Y_i - b_1 - b_2 X_i)^2$ 但是在WLS中,是最小化 $\sum \left(\frac{e_i}{\sigma_i}\right)^2 = \sum \left(\frac{Y_i - b_1 - b_2 X_i}{\sigma_i}\right)$,前提是σ_i 为已知。注意在WLS中我们是如何将大方差的观察值加以“压缩”的,因为误差方差越大,除数越大。

$$= B_1 \frac{1}{\sqrt{X_i}} + B_2 \sqrt{X_i} + v_i$$

其中 $v_i = u_i / \sqrt{X_i}$ 。也就是说，将模型(11-18)的两边同时除以 X_i 的平方根。方程(11-24)就是平方根变换(square root transformation)的一个例子。

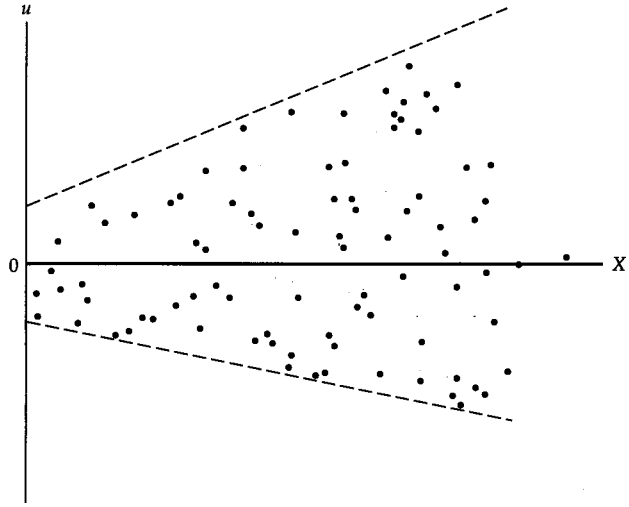


图11-8 误差与变量 X 成比例

根据等式(11-22)的思路，很容易证明变形后的回归方程的误差方差 v_i 是同方差的，因此，我们可以应用OLS法来估计方程(11-24)。事实上，这里我们正在使用WLS法(为什么？)[1]。有一点需要特别指出，为了估计方程(11-24)，我们必须使用过原点的回归进行估计。大多数标准回归软件包都能够做到这一点。(关于过原点的回归的详细讨论，参见习题6.15。)

例11.6 变换后的R&D回归方程。

对R&D数据，先前的检验表明可能存在异方差问题。根据图11-4，看起来误差与销售变量成比例。因此我们接受式(11-23)的假定以及估计方程(11-24)，得到如下结果：

$$\frac{\hat{Y}_i}{\sqrt{X_i}} = -246.68 \frac{1}{\sqrt{X_i}} + 0.0368 \sqrt{X_i}$$

$$\text{se}=(381.13) \quad (0.0071) \tag{11-25}$$

$$t=(-0.6472) \quad (5.1723) \qquad r^2=0.6258$$

将WLS结果与回归方程(11-3)中给出的OLS结果相比较，我们可以得出什么结论呢？

首先，读者也许会认为这两种结果无法比较，因为这两个模型中的被解释变量和解释变量是不相同的。但是，实际中的差异并没有那么明显，因为在求得方程(11-25)后，我们可将方程两边同乘以 $\sqrt{X_i}$ ，将其转化成：

$$\hat{Y}_i = -246.68 + 0.0368 X_i \tag{11-26}$$

该方程就可以与回归方程(11-3)直接进行比较了。两个方程的斜率系数相差很小。但是需注意的是，在方程(11-25)中，X 的系数比方程(11-3)相应的系数更为显著(为什

1 由于 $v_i = u_i / \sqrt{X_i}$，$v_i^2 = u_i^2 / X_i$，因此 $E(v_i^2) = \frac{E(u_i^2)}{X_i} = \sigma^2 \left(\frac{X_i}{X_i}\right) = \sigma^2$ ，也即同方差。注意变量 X 是非随机的。

(续)

么？)，这表明在回归方程(11-3)中，OLS法实际上高估了X系数的标准差。正如在前面指出的那样，当存在异方差时，标准差的OLS估计量是有偏的，而我们却无法知道它是偏高还是偏低。在这个例子中，它是偏高的，也就是说，它高估了标准差。就截距而言，它在两种情形下都是不显著的。

看来，式(11-23)的假定适合R&D一例，方程(11-25)所给出的WLS回归结果更具有说服力，因为我们已经考虑到了异方差问题。

问题是：如果模型中包括多个解释变量，情况又会怎样？在这种情况下，我们可以使用任何一个解释变量将模型变换成形如方程(11-24)那样的模式。比方说，我们可以根据图形找出合适的解释变量(参见习题11.7)。但是如果有多个解释变量都可以的话，又会怎样呢？在这种情况下，我们就不使用任何解释变量，而是利用估计的Y_i的均值$\hat{Y}_i$作为变换变量，因为$\hat{Y}_i$是解释变量X的线性组合。

情形2：误差方差与X_i^2成比例。如果所估计的残差呈现如图11-9所示的模式，则表明误差方差并不是与X线性相关，而是随着X的平方按比例增加。用符号表示，

$$E(u_i^2) = \sigma^2 X_i^2 \tag{11-27}$$

在这种情况下，我们需将方程的两边都除以X_i，而不是X_i的平方根，变换如下：

$$\frac{Y_i}{X_i} = B_1\left(\frac{1}{X_i}\right) + B_2 + \left(\frac{u_i}{X_i}\right)$$

$$= B_1\left(\frac{1}{X_i}\right) + B_2 + v_i \tag{11-28}$$

其中$v_i = u_i / X_i$

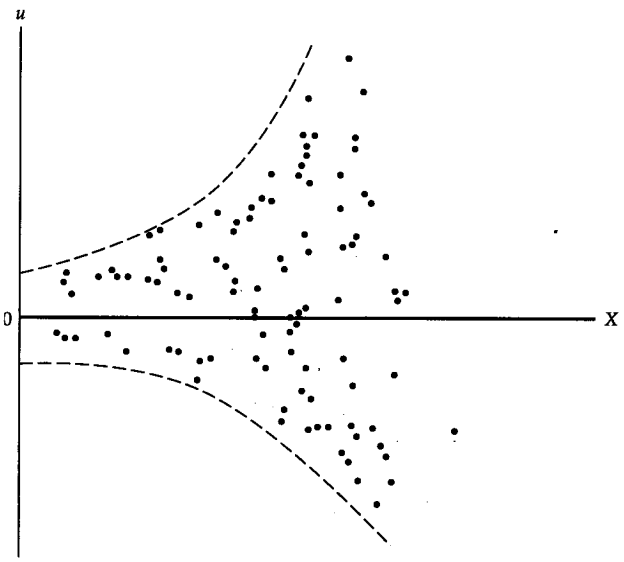


图11-9 误差方差与X^2成比例

按照前面的方法，读者可以很容易地验证方程(11-28)中的误差项v是同方差的。因此，用OLS法估计方程(11-28)，(这其实是WLS估计)可以得到最优线性无偏估计量(注意：古典线性回归中的其它假设依然不变)。

注意方程(11-28)中的一个值得关注的特性：原始方程中的斜率系数现在变成了“截距”，而原始方程中的截距现在则变成了“斜率”系数。但这一变化仅仅只是为了估计，一旦我们估计出方程(11-28)，将方程的两边同时乘以X_i，则又回到了原始模型。

11.4.3 重新设定模型

除了“推测”σ_i^2以外，有时我们重新设定总体回归函数(PRF)，即选择一个不同的函数形式(参见第8章)，这样也能够消除异方差。举例来说，如果我们不选择变量线性回归函数(LIV)的形式，而是用对数形式来估计模型，这样也常常能够消除异方差。即，如果我们估计：

$$\ln Y_i = B_1 + B_2 \ln X_i + u_i \tag{11-29}$$

在这个变换形式中的异方差问题也许就没那么严重，因为对数变换压缩了测定变量的尺度，从而把两个变量值间的10倍差异缩小为2倍差异。例如，90是9的10倍，但ln90 (=4.499 8)只有ln2 (=2.197 2)的两倍。

正如我们在第8章所看到的那样，对数线性或者双对数模型的一个显著优点是，斜率系数B_2度量的是Y对X的弹性，也即，X每变化百分之一时的Y变化的百分比。

在一个具体的例子中，是用线性变量模型还是对数线性模型要根据具体的理论以及其他一些因素来决定，我们将在第13章中详细讨论。但如果选择两者中的任何一个并没有太大的差别，并且在线性变量模型中异方差问题有比较严重时，则不妨试一试双对数模型。

例11.7 R&D对数线性模型

对于R&D数据，其对数线性回归模型如下：

$$\widehat{\ln Y_i} = -7.364\ 7 + 1.322\ 2 \ln X_i$$
$$\text{se}=(1.848\ 0) \quad (0.168\ 04) \tag{11-30}$$
$$t=(-3.985\ 2) \quad (7.868\ 7) \quad r^2=0.794\ 7$$

回归结果表明：在对数线性模型中，R&D支出与销售正相关，并且是显著相关的。估计的弹性系数表明，在其他条件保持不变的情况下，销售量每增加1%，平均而言，R&D支出将增加1.32%。

在习题11.9中，要求读者检证上述回归结果并检查是否存在异方差。如果回归方程(11-30)不存在异方差问题，则该模型显然优于变量线性模型，因为后者存在着异方差问题，需要对变量进行变换。

顺便指出：我们此前所讨论的消除异方差的变换方法也称为方差稳定变换(variance stabilizing transformations)。

为了作出我们所讨论的补救措施的结论，需要重申的是，以前所有讨论过的变换都是一种特别的变换；在不知道真实σ_i^2的确切信息的情况下，我们通常的做法是推测它会是什么样。究竟选择前面讨论的哪种变换，取决于问题的性质和异方差的严重程度。还要注意的是，有时候误差也许与模型中所包括的任何一个解释变量都不相关。相反，它可能与我们在建模型过程中可以考虑进去但实际上并未考虑进去的一个变量相关。在这种情形下，可以用这一变量来对模型加以变换。当然，按照逻辑，如果一个变量是属于模型的，则它在最开始就应该被包括进去。第13章对此会有更详细的论述。

11.5 White 异方差校正后的标准差和 *t* 统计量

前面我们已经指出，在存在异方差的情况下，OLS估计量尽管是无偏的，但却是无效的。

因此，按常规方法计算得到的估计量的标准差和 t 统计量都值得怀疑。White建立了一种估计方法，利用这种方法得到的标准差和回归系数考虑到了异方差的存在。因而，我们可以继续使用 t 检验和 F 检验，只不过这时的OLS估计量是渐近有效的，也即对大样本是有效的。[1]

下面讨论在存在异方差的情况下，用常规的 OLS法计算标准差和 t 统计量会导致什么样的后果，我们来看婴儿死亡一例 [回归方程(11-16)]。利用EVIEWS统计软件包，得到(White)异方差校正后的回归函数如下：

$$\hat{Y}_t = 111.781\ 3 - 0.004\ 2X_{2i} - 0.498\ 8X_{3i}$$
$$\text{se}=(38.726\ 4)\quad (0.001\ 4)\quad (0.424\ 2)$$
$$t=(2.886\ 4)\quad (-3.032\ 1)\quad (-1.154\ 6)\qquad R^2=0.492\ 3 \qquad (11\text{-}31)$$
$$p\text{值}=(0.010\ 3)\quad (0.007\ 5)\quad (0.264\ 2)\qquad \bar{R}^2=0.432\ 5$$

将未校正异方差回归方程(11-16)与回归方程(11-31)相比较，我们可以看到一些显著的变化。尽管两个回归方程中回归系数的值都是一样的，但标准差却不相同。在回归方程 (11-31)中，截距的标准差和变量 X_3 的标准差都显著变大，但变量 X_2 的标准差却略微减少。因此，我们无法再认为变量 X_3 对婴儿死亡率有显著影响。严格地说，虽然我们的样本观察值只有20个，也许还不够大到足于使用White方法，但它确实应证了如果怀疑存在异方差，那么，最好还是把它考虑进去。

11.6 实例

我们通过两个具体实例来结束本章的讨论，以此来说明在实际工作中异方差的重要性。

例11.8 规模经济或异方差

纽约股票交易所(NYSE)最初是极力反对对经纪佣金率放松管制的。事实上，在引入放松管制以前(1975年5月1日)，NYSE向股票交易委员会(SEC)提交了一份经济计量研究报告，认为在经纪行业中存在着规模经济，因此 (由垄断决定的)固定佣金率是公正的[2]。NYSE所提交的经济计量分析基本上是围绕着以下回归函数来进行的：[3]

$$\hat{Y}_i = 476\ 000 + 31.348X_i - (1.083\times 10^{-6})X_i^2$$
$$t=(2.98)\qquad (40.39)\qquad (-6.54)\qquad R^2=0.934 \qquad (11\text{-}32)$$

其中，Y=总成本，X=股票交易的数量。从模型(11-26)可以看出：总成本与交易量正相关。但是由于交易量的二次方项系数为负，并且是“统计显著的”，这意味着总成本是以一个递减的速率在增加。因此，NYSE认为在经纪行业中存在着规模经济，从而证明了NYSE的垄断地位是正当的。

然而美国司法部反托拉斯局却认为模型(11-32)中所声称的规模经济只是幻想，因为回归函数(11-32)存在着异方差问题。这是因为在估计成本函数时，NYSE并未考虑到样本中所包括的小公司与大公司的差别，也就是说，NYSE并没有考虑到规模因素。假设误差项与交易量成比例 (参见方程(11-23))，反托拉斯局重新估计了方程 (11-32)，

1 关于White异方差校正后的标准差的推导已超出了本书的范围。感兴趣的读者可以参考 Jack Johnston, John Dinardo,*Econometrics Methods*,4 th edition, New York，1997, Chap6.

2 一些经济学家提出了规模经济的论点来为一些行业的垄断正名，尤其是自然垄断行业 (例如，电力和煤气设备等)。

3 回归方程(11-32)和(11-33)中所给出的结果摘自H. Machael Mann，“The New York Exchange: A Cartel at the End of its Reign” in Almarin Phillips, *Promoting Competition in Regulated Industries*，Brookings Institution，Washtington D.C.,1975,p.324.

(续)

得到如下回归结果：[1]

$$\hat{Y}_i = 342\,000 + 25.57X_i + (4.34 \times 10^{-6})X_i^2$$

$$t=(32.3) \quad (7.07) \quad (0.503) \qquad (11\text{-}33)$$

从式(11-33)可以看出，二次项不仅是统计不显著的，而且其符号也发生了变化。[2] 因此，在经纪行业中并不存在规模经济，这就推翻了NYSE的垄断佣金结构的论点。

上述例子表明：方程(11-32)中所隐含的同方差假定的潜在危害有多大。设想一下，如果SEC接受了方程(11-32)的表面值并允许NYSE像1975年5月1日以前那样垄断地确定佣金率，那么情况会怎样!

例11.9 公路容量能力与经济增长

David A.Aschauer根据表11-3提供的结果证明了：拥有较好的陆地交通基础设施的经济将受益于更高的生产率和更快的人均收入增长。[3] 由于是对美国的48个州进行了调查研究，所以假定误差项可能不满足同方差假定"。[4]

表11-3 人均收入增长与高速公路运载能力

解释变量	OLS	WLS[1]	WLS[2]	WLS[3]
常量	-7.69	-7.94	-8.19	-7.62
	se=(1.06)	(1.08)	(1.09)	(1.08)
$\ln X_3$(以1960年计算)	-1.59	-1.64	-1.69	-1.58
	se=(0.18)	(0.19)	(0.19)	(0.18)
$\ln X_2$	0.30	0.30	0.31	0.30
	se=(0.06)	(0.06)	(0.06)	(0.06)
X_4	-0.009	−0.100	−0.011	-0.008
	se=(0.003)	(0.003)	(0.003)	(0.003)
D	-31.00	-32.00	-33.00	-31.00
	se=(0.08)	(0.08)	(0.08)	(0.08)
	R^2=0.67	0.49	0.46	0.73

注：1. 解释变量Y：从1960年到1980年的人均收入(按1972年的美元价计算)的年平均增长率。
2. X_2——以1960年为基年的人均收入水平(按1972年的美元价计算)
3. X_3——高速公路总距离(从1960年到1980年的平均水平)。
4. X_4——1982年低质量高速公路所占比例。
5. D——虚拟变量，中西部地区为1，其它地区为0。
6. WLS[1]——以$\ln X_2$的平方根为权数的加权最小二乘法。(参见方程(11-24))
7. WLS[2]——以$\ln X_2$水平为权数的加权最小二乘法。
8. WLS[3]——以$\ln X_2$水平为权数的加权最小二乘法。

资料来源：David. A. Aschauer，" Highway Capacity and Economic Growth, "*Economic Perspectives*, Federal Reserve Bank of Chicago，pp.14-23，September/October 1990，table 1，p.18，符号略作调整。

1 实际机制与估计方程(11-24)是一致的。一旦估计出这个方程，则将其乘以$\sqrt{X_i}$就可以得到原始方程，即方程(11-33)。

2 NYSE则认为反托拉斯局所使用的特定异方差假设是无效的。但其他替代形式的假设依然支持反托拉斯局的结论，即在经纪行业中并不存在规模经济。详细的讨论参见脚注19中Mann的文章。

3 这一例子和表11-3中所给出的统计结果摘自David. A. Aschauer，" Highway Capacity and Economic Growth, "*Economic Perspectives*, Federal Reserve Bank of Chicago，pp.14-23，September/October 1990.

4 异方差的拼写究竟是" heteroscedasticity "还是" heteroskedasticity "？应该是后面一个，但是前面的一个已经在文献中得到了广泛的承认以至于很少看到后面一个的拼写了。

(续)

令人高兴的是，在这个例子中，假定存在异方差是因为通过各种不同的方法校正异方差对OLS结果影响不大。从这个例子可以看出：如果存在着异方差，那么我们应该对这个问题作更深入的研究，而不是假定这个问题不存在。正如前面指出的那样，也如NYSE规模经济例子所说明的那样，异方差是一个潜在的严重的问题，决不能掉以轻心。在安全的一边犯错误会更好一些。

11.7 小结

古典线性回归模型的一个重要假设是扰动项 u_i 有相同的方差(也即同方差假定)。如果该假设不满足，则可能出现异方差问题。异方差并不破坏 OLS估计量的无偏性，但这些估计量不再是有效的。换言之，OLS估计量不再是最优线性无偏估计量。如果异方差 σ_i^2 已知，则用加权最小二乘法(WLS)可以得到最优线性无偏估计量。

忽视异方差的存在，而继续使用常用的 OLS法进行估计参数(这依然是无偏的)、建立置信区间和进行假设检验，则很可能得出错误的结论，正如 NYSE一例那样。这是因为所估计的标准差很可能是有偏的，从而导致了 t值也可能是有偏的。这样，在具体应用中确定是否存在异方差问题至关重要。对异方差检验有几种不同的方法，例如将估计残差对一个或者多个解释变量描图，Park检验，Glejser检验，秩相关检验等等。

如果通过一种或者多种检验表明存在着异方差问题，则需要补救措施。如果误差 σ_i^2 是已知的，则我们可以使用WLS法来获得最优线性无偏估计量。不幸的是，在实践中，很难获得有关真实的误差方差信息。所以我们不得不通过一些合理可行的假设来变换数据，使得变换后模型的误差项是同方差的。然后再使用将 OLS法，这实际上等同于WLS法。当然，要获得恰当的变换形式需要一定的技巧和经验。但是如果不进行变换，异方差的问题是无法得到解决的。

习题

11.1 异方差的含义是什么？它对下面各项有何影响？

(a) OLS估计量及其方差？ (b) 置信区间？ (c) 显著性t检验和F检验的使用？

11.2 以简单的理由来说明以下判断是对是错：

(a) 在存在异方差情况下，普通最小二乘法(OLS)估计量是有偏的和无效的。 (b) 如果存在异方差,通常用的t检验和F检验是无效的。 (c) 在存在异方差情况下,常用的OLS法总是高估了估计量的标准差。 (d) 如果从OLS回归中估计的残差呈现系统模式，则意味着数据中存在着异方差。

11.3 在如下回归中，你是否预期存在着异方差？

Y	X	样 本
(a) 公司利润	净财富	《财富》前500强
(b) 公司利润的对数	净财富的对数	《财富》前500强
(c) 道琼斯工业平均指数	时间	1960~1990年(年平均)
(d) 婴儿死亡率	人均收入	100个发达国家和发展中国家
(e) 通货膨胀率	货币增长率	美国、加拿大和15个拉美国家

11.4 从直观上解释，当存在异方差时，加权最小二乘法 (WLS)优于OLS法？

11.5 简要解释下列异方差诊断方法的逻辑关系：

(a) 图形法 (b) Park检验 (c) Glejser检验

11.6 在双变量总体回归函数中，假设误差方差有如下结构

$$E(u_i^2) = \sigma^2 X_i^4$$

你如何变换模型从而达到同方差的目的？你将如何估计变换后的模型？列出估计步骤。

11.7 考虑如下两个回归方程(根据1946到1975年美国数据)[1] (括号中给出的是标准差)：

$$\hat{C}_t = 26.19 + 0.624\,8\,\mathrm{GNP}_t - 0.439\,8 D_t$$

$$\mathrm{se} = (2.73) \quad (0.006\,0) \quad (0.073\,6) \qquad R^2 = 0.999$$

$$\left(\frac{\hat{C}}{\mathrm{GNP}}\right)_t = 25.92\frac{1}{\mathrm{GNP}_t} + 0.624\,6 - 0.4315\frac{D}{\mathrm{GNP}_t}$$

$$\mathrm{se} = (2.22) \quad (0.006\,8) \quad (0.059\,7) \qquad R^2 = 0.875$$

式中 C —— 总私人消费支出

GNP —— 国民生产总值

D —— 国防支出

T —— 时间

Hanushek和Jackson研究的目的是确定国防支出对经济中其它支出的影响。

(a) 将第一个方程变换成第二个方程的原因是什么？ (b) 如果变换的目的是为了消除或者减弱异方差，那么对误差项要作哪些假设？ (c) 如果存在异方差，是否已成功地消除异方差？你是如何知道的？ (d) 变换过回归方程是否一定要通过原点？为什么？ (e) 能否将两个回归方程中的R^2加以比较？为什么？

11.8 在研究人口密度对离中心商业区距离的回归函数中， Maddala根据1970年巴尔蒂摩地区39个人口普查区的有关数据得到如下回归结果[2]：

$$\widehat{\ln Y_i} = 10.093 - 0.239 X_i$$

$$t = (54.7) \qquad (-12.28) \qquad R^2 = 0.803$$

$$\frac{\widehat{\ln Y_i}}{\sqrt{X_i}} = 9.932\frac{1}{\sqrt{X_i}} - 0.225\,8\sqrt{X_i}$$

$$t = (47.87) \qquad (-15.10)$$

式中 Y —— 普查区的人口密度

X —— 离中心商业区的距离(英里)

(a) 如果存在异方差，作者在其数据中作了哪些假设？ (b) 从变换后的(WLS)回归函数中，你如何知道异方差已被消除或减弱了？ (c) 你如何解释回归结果？它是否有经济意义？

11.9 参考表11-2给出的R&D数据。回归方程(11-30)给出了对数形式的R&D和销售的回归结果。

(a) 根据表11-2提供的数据，验证这个回归结果。 (b) 分别将残差的绝对值和残差平方值对销售量描图。是否表明存在着异方差？ (c) 对回归的残差进行Park检验和Glejser检验。你得出什么结论？ (d) 如果在双对数模型中发现了异方差，你会选择用哪种WLS变换来消除它？ (e) 如果对线性回归函数(11-3)，有证据表明存在着异方差。而在对数 -对数模型中没有证据表明存在异方差，那么，你将选择哪个模型，为什么？ (f) 能够比较两个回归方程的R^2吗？为什么？

11.10 参考表11-2中提供的R&D数据，现考虑如下回归方程：

$$\mathrm{R\&D}_i = A_1 + A_2 \mathrm{profits}_i + u_i$$

$$\ln \mathrm{R\&D}_i = B_1 + B_2 \ln \mathrm{profits}_i + u_i$$

---

1 这些结果取自Eric A.Hanushek，John E. Jackson，Statistical Methods for Social Scientists，Academic，New York，1977，p.160。

2 G.S.Maddala，*Introduction to Econometrics*，Macmillan，New York，1988，p.175-177.

(a) 估计两个回归方程。 (b) 求每个回归方程的残差的绝对值和残差平方，并分别将它们对解释变量描图。你是否诊断到存在异方差？ (c) 用Glejser检验和Park检验验证你在(b)中定性分析得到的结论。 (d) 如果有迹象表明存在着异方差，你如何变换数据以削减异方差？给出必要的计算步骤。

11.11 考虑图11-10，该图描绘了从1974年到1985年期间一些国家的国内生产总值(GDP)增长的百分率与投资/GDP率之间的关系。[1] 这些国家分为不同的三类——实际利率为正的国家，实际利率为适中负值的国家，以及实际利率为较大负值的国家。

(a) 建立合适模型解释GDP增长的百分率与投资/GDP百分比之间的关系。 (b) 从图中你能发现数据中存在异方差吗？你是如何正规地检测异方差的存在的。 (c) 若怀疑存在异方差，那么，你如何变换回归函数以消除异方差？ (d) 假设你通过加入虚拟变量来扩展模型以便考虑到这三类国家“质”的不同。写出该回归方程。如果根据数据能够估计这一扩展模型，你会预期扩展的模型中存在异方差吗？为什么？

11.12 1964年，对9 966名经济学家的调查数据如下：

年龄/岁	中值工资/美元	年龄/岁	中值工资/美元
20~24	7 800	50~54	15 000
25~29	8 400	55~59	15 000
30~34	9 700	60~64	15 000
35~39	11 500	65~69	14 500
40~44	13 000	70+	12 000
45~49	14 800		

资料来源：“The Structure of Economists'Employment and Salaries,”Committee on the National Science Foundation Report on the Economics Profession, *American Economics Review*, vol.55,no.4, December 1965, p.36.

(a) 建立适当的模型解释平均工资与年龄间的关系。为了分析的方便，假设中值工资是年龄区间中点的工资。 (b) 假设误差与年龄成比例，变换数据求得WLS回归方程。 (c) 现假设误差与年龄的平方成比例，求WLS回归方程。 (d) 哪一个假设看来更可行？

11.13 异方差的Sperman秩相关检验。该检验包括以下步骤，可以用R&D回归方程(11-3)加以说明。

(a) 从回归方程(11-3)中求得残差e_i。 (b) 求残差的绝对值，$|e_i|$。 (c) 将销售X_i和$|e_i|$或者按降序(从高到低)或者升序(从低到高)排列。 (d) 对于每个观察值，取两列间的差，称之为d_i。 (e) 计算Sperarman秩相关系数r_S，定义为

$$r_S = 1 - 6\left[\frac{\sum d_i^2}{n(n^2-1)}\right]$$

其中，n=样本中观察值的个数。

如果在$|e_i|$和X_i之间存在系统关系，则二者之间的秩相关系数将是统计显著的，在这种情况有理由怀疑存在异方差。

给定零假设：真实总体秩相关系数为零，且$n>8$，则可以证明

$$\frac{r_S\sqrt{(n-2)}}{\sqrt{1-r_S^2}} \sim t_{n-2}$$

服从自由度为$(n-2)$的学生t分布。

1 参见World Development Report,1989,the World Bank,Oxford University Press,New York,p.33.

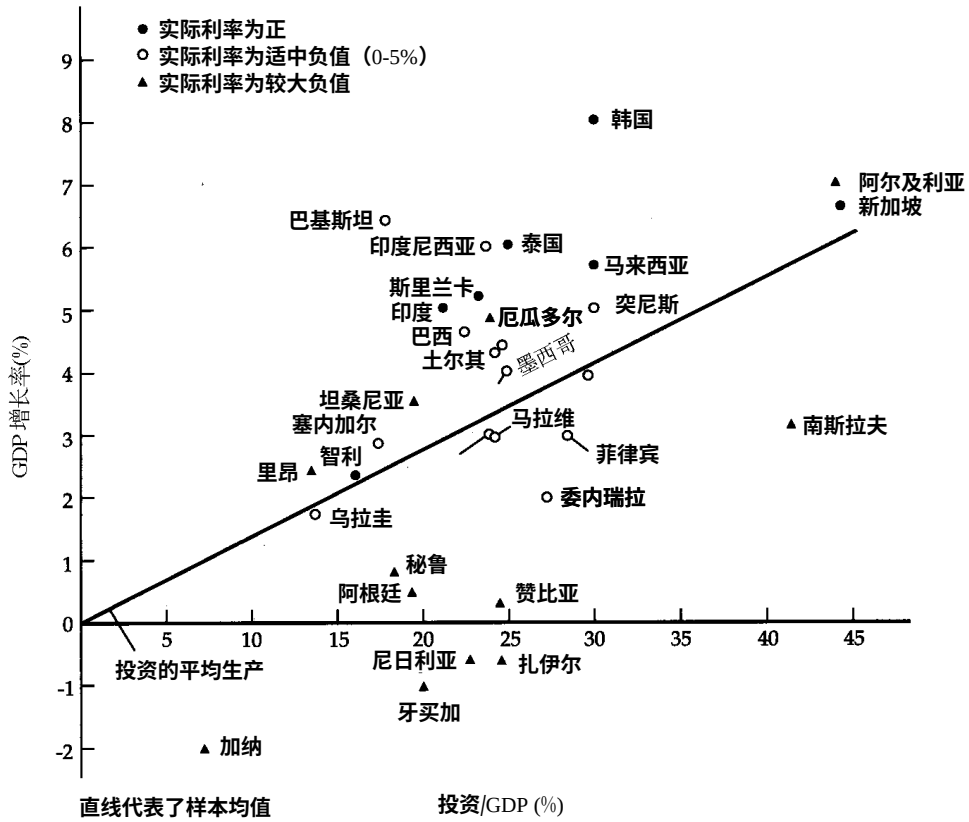


图11-10 1974到1985年33个发展中国家的实际利率，投资，生产率和增长率

因此，在实际应用中，根据t检验，若秩相关系数是显著的，则可以接受假设：存在异方差。用这个方法检验本章的R&D数据，确认数据存在着异方差。

11.14 加权最小二乘法。考虑表11-4中的数据。

(a) 估计OLS回归方程

$$Y_i=B_1+B_2X_i+u_i$$

表11-4 美国制造业平均赔偿与就业规模所决定的生产率之间的关系

就业规模(平均就业人数)	平均赔偿Y/美元	平均生产率X/美元	赔偿的标准方差σ_i/美元
1~4	3 396	9 335	744
5~9	3 787	8 584	851
10~19	4 013	7 962	728
20~49	4 104	8 275	805
50~99	4 146	8 389	930
100~249	4,241	9 418	1 081
250~499	4 387	9 795	1 243
500~999	4 538	10 281	1 308
1 000~2 499	4 843	11 750	1 112

数据来源：*The census of Manufacturing*，U.S Department of Commerce，1958.(表中的数字由D.N.Gujarati计算)

(b) 估计WLS

$$\frac{Y_i}{\sigma_i}=B_1\frac{1}{\sigma_i}+B_2\frac{X_i}{\sigma_i}+\frac{u_i}{\sigma_i}$$

(确保运行的WLS通过原点。)计算两个回归方程的结果。你认为哪个回归方程更好？为什么？

11.15 证明方程(11-28)中的误差项v_i是同方差的。

11.16 在平均工资对就业人数的回归分析中(包括30个公司的一随机样本)，得到如下回归结果：[1]

$$\hat{W} = 7.5 + 0.009N \tag{1}$$

$$t = \text{N.A.} \qquad (16.10) \qquad R^2 = 0.90$$

$$\frac{\hat{W}}{N} = 0.008 + 7.8\frac{1}{N}$$

$$t = (14.43) \quad (76.58) \qquad R^2 = 0.99 \tag{2}$$

(a) 你如何来解释这两个回归方程？ (b) 从方程(1)到(2)，作者作了哪些假设？他是否担心存在异方差？ (c) 你能将两个回归方程中的斜率和截距联系起来吗？ (d) 你能比较两个回归方程中的R^2吗？为什么？

11.17 根据总成本函数(11-30)，你能推导出平均成本函数、边际成本函数吗？如果方程(11-33)是真实的(异方差调整之后)总成本函数，你又如何推导出相应的平均成本函数和边际成本函数？解释两个模型之间的差异。

11.18 表11-5给出了20个国家五项社会经济指标的有关数据，分成四个人均收入等级：低收入(每年500美元以下)，中等偏低收入(年收入在500到2 200美元之间)，中等偏上收入(年收入在2300到5 500美元之间)，以及高收入(年收入超过5 500美元)。表中前五个国家属于低收入国家，其次五个国家属于中等偏低收入国家，以此类推。

(a) 假设你想在模型(11-16)中加入补充解释变量人口增长率X_4和每日卡路里(calorie)吸收量X_5。根据推理，你预期这些变量对婴儿的死亡率Y有什么影响？ (b) 估计上述回归方程并检验你的预期是否正确。 (c) 如果在上述方程中遇到了多重共线性问题，你怎么办？可以采取任何你认为正确措施。

11.19 按照(White)回归方程(11-14)的过程校正了异方差之后，方程(11-16)的回归结果如下：(注：为了节省空间，我们只给出了t统计量和它们的ρ值。这些结果是通过MINITAB 软件包获得的)

表11-5 20个国家的婴儿死亡率

国 家	IMOR①	PCGNP②	PEDU③	POPGROWTH④	CSPC⑤
坦桑尼亚	104	160	66	3.5	2 192
尼泊尔	126	180	82	2.6	2 052
马里	168	230	23	2.4	2 073
尼日利亚	103	290	77	3.3	2 146
加纳	88	400	71	3.4	1 759
菲律宾	44	630	18	2.5	2 372
科地瓦尔	53	770	22	4.0	2 562
危地马拉	57	900	77	2.9	2 307
土耳其	75	1 280	117	2.3	3 229
马来西亚	23	1 940	102	2.6	2 730
阿尔及利亚	72	2 360	96	3.1	2 715
乌拉圭	23	2 470	110	0.6	2 648
韩国	24	3 600	101	1.2	2 907
希腊	12	4 800	104	0.5	3 688

1 参见Dominick Salvatore，*Managerial Economics*，McGraw-Hill, New York，1989，p.157.

(续)

国 家	IMOR①	PCGNP②	PEDU③	POPGROWTH④	CSPC⑤
委内瑞拉	25	3 250	107	2.8	2 494
西班牙	9	7 740	113	0.5	7 740
以色列	11	8 650	95	1.7	3 061
澳大利亚	9	12 340	106	1.4	3 326
英国	9	12 810	106	0.2	3 256
美国	10	19 840	100	1.0	3 645

① IMOR —— 婴儿死亡率(每千个出生婴儿中)，1988。
② PCGNP —— 人均GNP(1988年美元)。
③ PEDU —— 初等教育入学年龄集团所占百分率，1987。
④ POPGROWTH —— 人口增长率，1980～1988年平均值。
⑤ CSPC —— 人均每日卡路里供应量，1986。

资料来源：World bank, *world Development Report*, 1990.

$$\hat{Y}_i = 498\ 7 - 0.4718X_{2i} - 0.844\ 2X_{3i} + 0.000\ 010\ 2X_{2i}^2 + 0.443\ 5X_{3i}^2 + 0.002\ 6X_{2i}X_{3i}$$

$$t=(4.86)\quad(-0.59)\quad(-2.45)\quad(1.27)\quad(1.62)\quad(0.35)$$

$$p\text{值}=(0.000)\quad(0.566)\quad(0.028)\quad(0.223)\quad(0.127)\quad(0.731)\quad R^2=0.649$$

(a) 你如何解释上述回归方程？ (b) 回归方程(11-17)给出了White异方差检验的结果，该结果表明回归方程(11-16)存在着异方差问题。你将对数据采取什么变换以消除异方差？ (提示：参考第11.4节)给出必要的计算步骤。 (c) 你如何确定(b)中所选的模型是否存在异方差问题？

11.20 (a)利用表11-5提供的数据建立一个多元回归模型用以解释表中所示的20个国家的每日卡路里吸入量。 (b)该模型是否存在着异方差问题？给出必要的检验。 (c) 如果存在着异方差，求White异方差校正后的标准差和 t 统计量(看看你的软件是否能做到这一点)，与(a)中的结果相比较并作出评论。

第12章

自相关

在上一章中，我们对放松古典线性回归模型(CLRM)的假设之一——同方差假定的后果进行了考察。本章，我们考虑放松CLRM另一假设——总体回归函数(PRF)的扰动项u_i无序列相关(serial correlation)或自相关(autocorrelation)。我们曾在第6章中对这一假设作过简单讨论，本章将深入讨论这一问题，以寻求以下问题的答案：

(1) 自相关有什么性质？

(2) 自相关的理论与实际结果是什么？

(3) 由于非自相关假设与不可观察的扰动项u_i有关，那么，如何判断在给定情况下存在自相关？简言之，在实际中，如何诊断自相关？

(4) 如果发现自相关的后果比较严重，如何采取措施加以补救？

本章对自相关的讨论在许多方面与上一章的异方差问题相类似。在存在异方差和自相关的情况下，普通最小二乘法估计量，尽管是线性的和无偏的，但却不是有效的。也即，它们都不是最优线性无偏估计量。

本章重点讨论自相关问题，在此假设CLRM中的其他假定保持不变。

12.1 自相关的性质

自相关一词可以定义为，“在时间(如在时间序列数据中)或者空间(如在横截面数据中)按顺序所列观察值序列的各成员间存在着相关”。[1]

正如异方差的产生通常是与横截面数据有关，自相关问题通常是与时间序列数据有关(也即，数据按照时间顺序排列)。根据上述定义，在横截面数据中也可能产生自相关问题，这种情况下，称为空间相关(spatial correlation)。

在古典线性回归模型中假定在扰动项u_i中不存在自相关。用符号表示为：

$$E(u_i u_j)=0 \qquad i \neq j \tag{12-1}$$

也就是说，两个不同误差项u_i和u_j的乘积的期望为零。[2]简单地说，古典模式假定任一观察值的扰动项不受其他观察值的扰动项的影响。例如，在讨论产出对劳动和资本投入回归(也即生产函数)的季度时间序列数据时，譬如说，某一季度工人罢工影响了产出，但却没有理由设

1 Maurice G.Kendall和William R.Buckland, *A Dictionary of Statistical Terms*，Hafner，New York，1971，p.8.

2 如果$i=j$，方程(12-1)就变成了$E(u_j^2)$，也即u_j的方差，根据同方差假定，等于σ^2。

想这一“中断”会持续到下一个季度。换言之，如果本季度产出降低，但并不意味着下一季度产出也下降。类似地，在分析家庭消费支出与家庭收入的横截面数据时，一个家庭收入增加对其消费支出的影响并不会影响另一个家庭的消费支出。

但是如果存在这种依赖关系，便产生了自相关问题。用符号表示，

$$E(u_i u_j) \neq 0 \qquad i \neq j \tag{12-2}$$

在这种情形下，本季度由于罢工引起生产“中断”可能对下一季度的产出有影响(事实上，产出可能增加以弥补上一季度生产的不足)；或者，某一家庭消费支出的增加可能马上影响不愿比别人逊色的另一家庭消费支出的增加(这就是空间相关的例子)。

图12-1给出了自相关和非自相关的类型。在纵轴上，同时给出了 u_i(总体扰动项)及相应的(样本扰动项) e_i，因为，与异方差情形相同，我们无法观察到前者而只能通过后者来推导前者的行为。图12-1a到d表明了u中存在可辨别的模式，而图12-1e则表明不存在系统模式，表示支持古典线性回归模型(12-1)关于无自相关的假定。

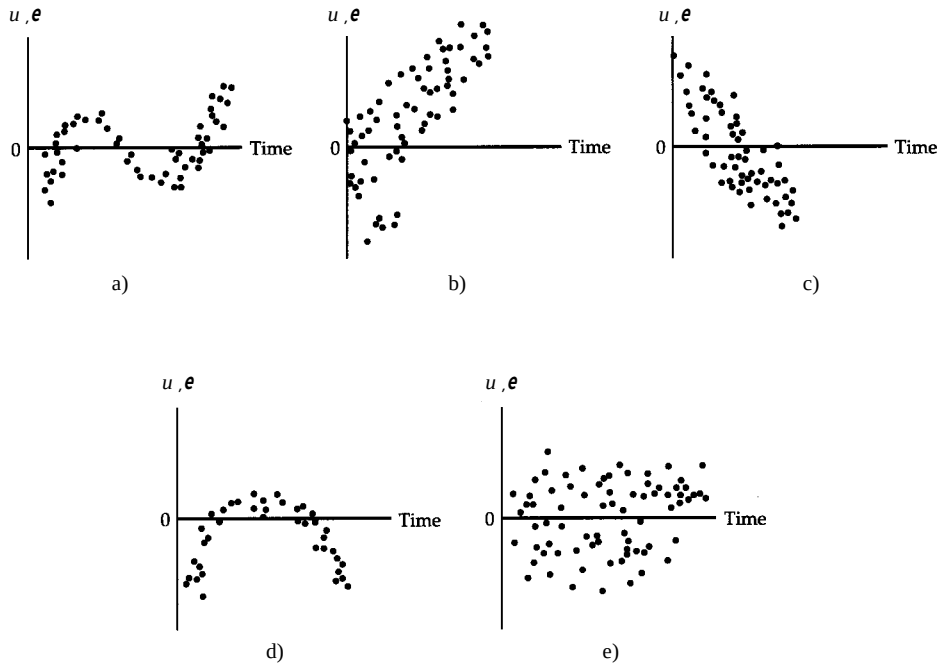


图12-1 自相关的模型

大家很自然会提出这样的问题：为什么会产生自相关呢？这里面有许多原因，其中几个如下：

12.1.1 惯性

大多数经济时间序列的一个显著特征就是惯性(inertia)或者说是迟缓性(sluggishness)。众所周知，时间序列，例如国民生产总值、就业、货币供给、价格指数等等，都呈现商业循环(在经济活动中重复发生或者自我维持波动)。当经济恢复开始时，由萧条的底部开始，大多数的经济序列向上移动。在向上移动的过程中，序列某一时点的值会大于其前期值。这里有一种“动力”存在，继续向上，直到某些事件发生(例如税收的增加或者利率的提高或者两者同时增加)才使序列移动减慢下来。因此，在涉及时间系列数据的回归方程中，连续的观察值之间很可能是相关的。

12.1.2 模型设定误差

有时候，形如图12-1a到d的自相关的发生并不是因为连续观察值之间相关，而是由于回归模型没有“正确的”设定。模型的不恰当设定意味着或是由于本应包括在模型中的重要变量未包括进模型中(这是过低设定的情形)，或是模型选择了错误的函数形式——本应该使用对数线性模型但却用了线性变量模型(LIV)。如果发生这样的模型设定误差(model specification errors)，则从不正确的模型中得到的残差将会呈现系统模式。一个简单的检验方法是将略去的变量包括到模型中，判定残差是否仍然呈现系统模式。如果它们并不存在着显著模式，那么系列相关可能是由于模型设定的错误。下一章会就这个问题进行详细的讨论(另见习题 12.13和12.18)。

12.1.3 蛛网现象

许多农产品的供给都呈现出所谓的蛛网现象(the cobweb phenomenon)，即供给对价格的反应滞后了一个时期，因为供给决策的实现需要一定的时间——(有一孕育期)。因此，种植谷物的农民本年度的计划是受上一年度价格的影响，所以他们的供给函数为：

$$\text{供给}_t = B_1 + B_2 P_{t-1} + u_t \tag{12-3}$$

假设在t时期末，价格P_t竟然低于P_{t-1}，于是，在$(t+1)$末时期，农户们决定比t时期少生产一些。显然，在这种情形下，扰动项u_t并不设想它是随机的，因为，如果农民在第t年生产多了，则他们很可能会在第$(t+1)$年少生产一些，这样下去，就会形成蛛网模式。

12.1.4 数据加工

在实证分析中，通常原始数据是要经过加工的。例如，在季度数据的时间序列回归中，数据通常是通过月度数据推导而来的，即将3个月的数据简单加总并除以3。这样平均的结果，消除了月度数据的波动性，而这种“平滑”过程本身就可能导致扰动项的系统模式，从而引入自相关。[1]

在继续下面的讨论之前，需要指出的是：虽然大多数经济时间序列都因在一个时期内或者上升或者下降而表现出正的自相关，而不像图12-2b那样表现为一上一下的恒常运动，但是，自相关可能是负的，也可能是正的。

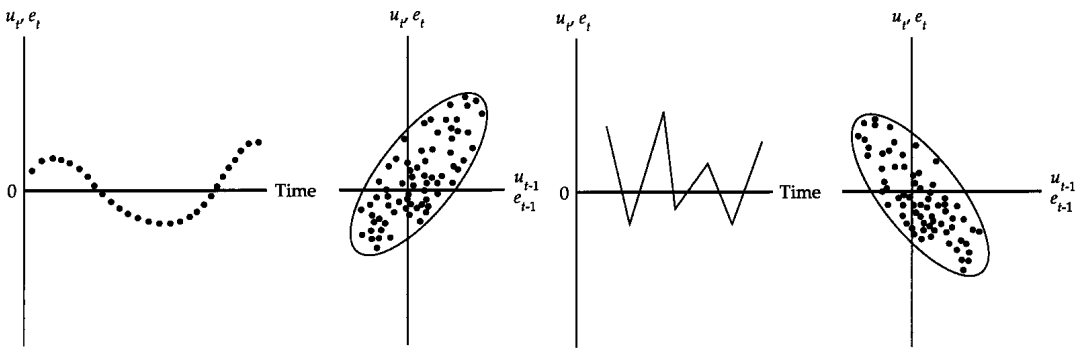


图 12-2
a)正的自相关 b)负的自相关

1 需要指出的是，有的时候我们用求平均或者其他数据处理方法是因为周数据或者月度数据存在着许多测量误差。因此，求平均的过程能够得到更为精确的估计值。但是这一过程的副作用就是它可能会引入自相关。

12.2 自相关的后果

假设误差项呈现如图12-1a到b或图12-2所示的其中一种模式。那么会产生什么后果呢？换言之，放松假设(12-1)，使用OLS法会导致什么后果呢？[1]

(1) 最小二乘估计量仍然是线性的和无偏的。

(2) 但却不是有效的。简言之，通常所用的普通最小二乘(OLS)估计量并不是最优线性无偏估计量(BLUE)。

(3) OLS估计量的方差是有偏的。有时候，用来计算方差和OLS估计量标准差的公式会严重低估真实的方差和标准差，从而导致 t 值变大。这会使得从表面上看某个系数显著不为零，但是事实却并非如此。

(4) 因此，通常所用的 t 检验和 F 检验一般来说是不可靠的。

(5) 计算得到的误差方差，$\hat{\sigma}^2$ =RSS/d.f.(残差平方和/自由度)，是真实 σ^2 的有偏估计量，在有些情形下，它很可能是低估了真实的 σ^2。

(6) 因此，通常计算的 R^2 不能测度真实 R^2。

(7) 通常计算的预测的方差和标准差可能也是无效的。

正如读者所看到的那样，自相关产生的后果与异方差产生的后果很相似，在实践中，这些后果同样也是严重的。因此，与异方差情况相同，在具体的应用中，我们必须确定我们是否存在自相关问题。

12.3 自相关的诊断

说到自相关的诊断(detecting autocorrelation)，我们遇到了与在异方差情形时一样的两难选择。在这里，我们并不知道误差的 σ^2 的真实值，因为真实的 u_i 是无法观察的。还有一个问题就是，我们不但不知道真实的 u_i 是什么，而且如果它们是相关的，我们也不知道其产生机制：我们仅仅有它们的替代物 e_t。因此，与异方差情形一样，我们不得不根据从标准OLS法中得到的 e_t 来“了解”自相关的存在与否。带着这一告诫，我们介绍自相关的几种诊断方法，我们用第6章中遇到的一个例子来加以说明。

例12.1 美国进口支出函数

在回归方程(6-52)中，我们给出了1968年到1982年期间与个人可支配收入(PDI)相关的美国进口支出函数的结果，所有的数据都是以1982年美元价计算(10亿美元)。在表12-1中，我们给出了残差，残差的一期滞后值以及残差的符号。我们用表中提供的数据说明所要讨论的各种不同的诊断检验方法。

表12-1 从回归方程(6-52)中所获得的残差

	e_t	e_{t-1}	$D=e_t-e_{t-1}$	D^2	e_t^2	Signe_t
	(1)	(2)	(3)	(4)	(5)	(6)
e_1	16.364 2	—	—	—	267.787 3	+
e_2	13.370 5	16.364 2	−2.993 7	8.962 3	178.770 1	+
e_3	2.921 2	13.370 5	−10.449 3	109.187 7	8.533 4	+
e_4	3.433 8	2.921 2	0.512 6	0.262 7	11.790 7	+

1 证明可参见：Damodar N.Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，1995，Chap.12.

(续)

	e_t	e_{t-1}	$D=e_t-e_{t-1}$	D^2	e_t^2	Signe_t
	(1)	(2)	(3)	(4)	(5)	(6)
e_5	11.012 8	3.433 8	7.579 0	57.441 9	121.281 8	+
e_6	9.354 8	11.012 8	−1.658 0	2.748 9	87.512 5	+
e_7	7.712 3	9.354 8	−1.642 5	2.697 8	59.479 5	+
e_8	−24.721 8	7.712 3	−32.434 0	105 1.967 0	611.165 0	−
e_9	0.283 7	−24.721 8	25.005 5	625.273 8	0.080 5	+
e_{10}	13.696 6	0.283 7	13.412 8	179.904 0	187.595 6	+
e_{11}	3.677 2	13.696 6	−10.019 3	100.386 5	13.522 2	+
e_{12}	−3.607 2	3.677 2	−7.284 4	53.063 2	13.011 9	−
e_{13}	−28.324 1	−3.607 2	−24.716 9	610.924 8	802.254 5	−
e_{14}	−31.635 5	−28.324 1	−3.311 4	10.965 6	100 0.807 0	−
e_{15}	−43.999 0	−31.635 5	−12.363 5	152.855 7	193 5.913	−
e_{16}	−28.563 3	−43.999 0	15.435 7	238.261 5	815.861 6	−
e_{17}	6.519 3	−28.563 3	35.435 7	123 0.790 0	42.348 7	+
e_{18}	5.417 4	6.519 3	−1.101 9	1.214 1	29.348 7	+
e_{19}	25.760 3	5.417 4	20.342 8	413.830 2	663.590 9	+
e_{20}	41.326 8	25.760 3	15.566 5	242.316 0	170 7.901 0	+
总计	0			509 3.140 0	855 8.8	

注：数据四舍五入至小数点后四位数。

12.3.1 图形法

与异方差情形相同，通过对OLS残差e的直观检验判断误差项u中是否存在自相关。有几种不同的检验残差方法。我们可以将残差对时间描图，如图 12-3所示，该图描绘了表12-1的进口支出函数的残差。这种图形称为时间序列图(time-sequence plot)(另见表6-3)。

对于图12-3的检验表明：残差e并不是如图12-1e所示那样随机分布。事实上，它们呈现一种明显的变动行为——开始，它们通常是正的，接着变成负值，再又变成正值。如果将t时间的残差与滞后一期的残差值描图，也即将表 12-1中第1列的e_t对第2列中的e_{t-1}描图(如图12-4所示)，则会看得更清楚。该图表明存在着正的自相关：大多数残差都分布在第一象限(西北方向)和第三象限(东南方向)。

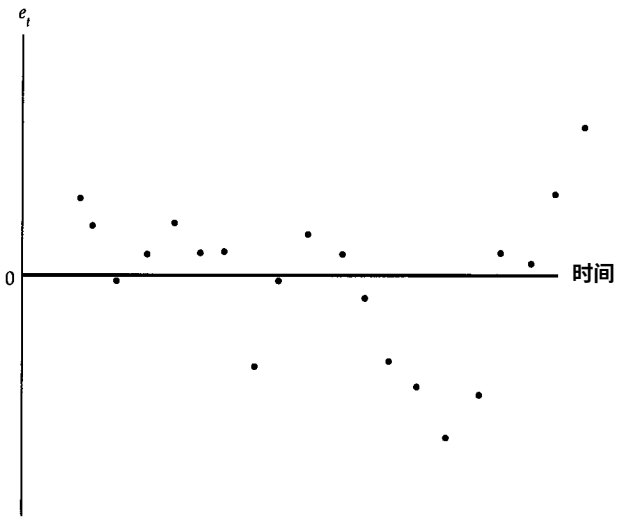


图12-3 回归方程(6-52)的残差

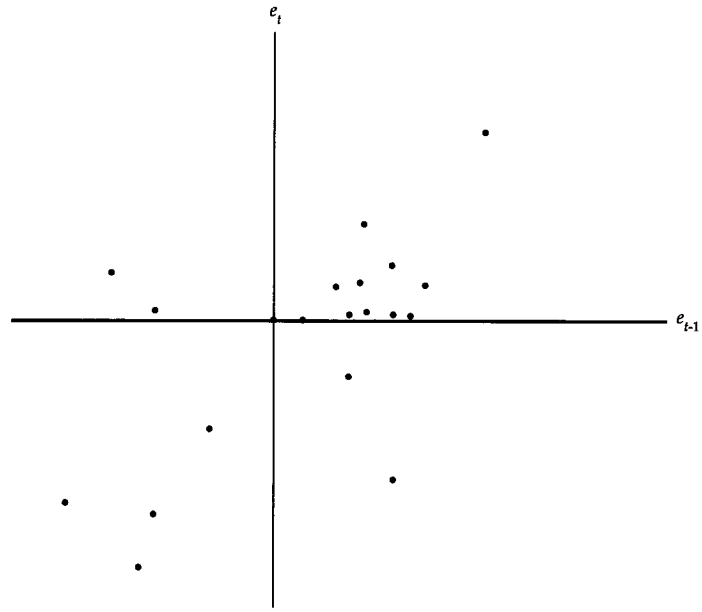


图12-4 回归方程(6-52)的残差e_t和e_{t-1}

12.3.2 游程检验

既然图形已表明在进口支出数据中可能存在着 (正的)自相关，我们可以利用一些正式的方法来补充定性的猜测。图12-3和表12-1给出的残差检验表明一些正的残差后面是一些负的残差，再又有一些正的残差。如果残差果真是随机的，那么我们能够观察到这种模式吗？直观地看是不可能的。我们可以用游程检验(runs test)，一种非参数检验方法来加以验证。[1]

为了说明这一检验，我们记录下残差的符号 (+或者－)，表12-1的第6列给出了不同残差的符号。我们用下面的方式写出这些符号：

$$(+\ +\ +\ +\ +\ +\ +)(-)(+\ +\ +)(-\ -\ -\ -\ -)(+\ +\ +\ +) \qquad (12\text{-}4)$$

这样，有7个正的残差跟着一个负的残差，再跟着 3个正的残差，再跟着 5个负的残差，再跟着4个正的残差，总共是 20个残差。我们现在将“ 游程(run) ”定义为同一符号或属性(例如+或-)的一个不间断历程。进一步定义“ 游程 ”的长度(length of the run)为游程中正负交替的个数。在式(12-4)所示的序列中共有 5个游程 —— 一个7个正值的游程(也就是说长度为 7)，一个仅为1个负值的游程(也就是说长度为1)，一个3个正值的游程(也就是说长度为3)，一个5个负值的游程(也就是说长度为 5)以及一个 4个正值的游程 (也就是说长度为 4)(为了有更好的观察效果，我们已用括号区分开不同的游程)。

通过考察一个随机观察次序中这些游程的表现形式，我们可以得出随机游程检验。我们要问的问题是：与预期相比，从这 20个观察值中所观察到的 5个游程是太多了还是太少了？如果游程太多了，它意味着 e在频繁地变换着符号，表明存在负的序列相关 (比较图12-2b)。类似地，如果游程太少，则意味着正的自相关，如图 12-2a所示。

令N—— 观察值的总个数$(=N_1+N_2)$；

N_1—— ＋号(也就是正的残差)的个数；

N_2—— －号(也就是负的残差)的个数；

1 在非参数检验中，我们对观察值的(概率)分布不作任何假设。

k——游程个数。

接着，在残差是独立的假设下，史威德 (Swed)和艾森哈特(Eisenhart)建立了一些特殊的表格，这些表格给出了在N个观察值的随机次序下预期游程的临界值。这些表格在附录 A，表A-6中给出。

Swed-Eisenhart临界的游程临界值 (Swed-Eisenhart critical runs test)。我们用上面的这个例子说明这些表格。这里，有 N=20，N_1=14(14个正值)，N_2=6(6个负值)以及k=5。对于N_1=14，N_2=6，5%的临界的游程值为 5。现在如果实际游程个数等于或者小于 5，则我们可以拒绝这一假设，即式 (12-4)中e是随机的。在我们的例子中，实际的游程个数是 5。因此，我们拒绝零假设：e是随机的。换言之，我们可以下结论说，我们的进口支出数据中存在着自相关问题。

例12.2 抵押债务

我们再来看一个例子。参考抵押债务回归方程 (7-32)，附录7A.4给出了回归结果。从这一结果中，你可以验证，n=16，n_1(+残差数)=7，n_2(−残差数)=9，游程个数=4。现在从Swed-Eisenhart表格中，我们注意到，5%的临界值为4和14，意味着如果观察到的数等于或小于4或等于或大于14，则表明存在着自相关。也就是说，我们的抵押债务回归方程存在着自相关问题。

注意Swed-Eisenhart表格最多为40个观察值，20个正值及20个负值。但如果实际样本更大的话，我们就无法使用这些表格了。但是在这种情形下，游程的分布可以由正态分布来近似。这一情形下的详细内容可以查阅有关参考书。[1]

12.3.3 杜宾－瓦尔森d检验[2]

诊断自相关最为著名的检验是由杜宾和瓦尔森提出的，杜宾－瓦尔森 d统计量，定义为：

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \tag{12-5}$$

它是逐次残差差的平方和对残差平方和的比值。注意在 d统计量分子的容量为 $(n-1)$，因为在求逐次差时失去一个观察值。

d统计量的一个最大优点是简单易行，它以 OLS残差为基础，而许多回归软件包都可以对残差进行计算。通常，统计结果在给出 R^2，校正的R^2($\bar{R}^2$)，t比例，F比例等的同时，也给出杜宾－瓦尔森d统计量。(参见表6-4所给出的SHAZAM结果)。

对进口支出一例，利用表 12-1所给的数据，可以很容易地计算出 d统计量。首先，将该表第1列中的残差减去第 2列中残差的滞后值，再将差的平方求和，然后除以表中第 5列残差的平方和。计算 d所需的必要原始数据都在表 12-1中给出。当然，这一切现在已由计算机来完成。在我们的这个例子中，计算的 d值为：

$$d = \frac{5093.14}{8\,558.8} = 0.595\,1 \tag{12-6}$$

在继续说明如何利用 d值判定自相关存在与否之前，注意作为 d统计量最基础的一些假设是

1 详细内容参见，Damodar N.Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，1995，p.419-420.

2 J. Dubin，G.S. Watson，“Testing for Serial Correlation in Least-Squares Regression,”*Biometrika*，vol.38，1951，p.159-177.

非常重要的：

(1) 回归模型包括一个截距项。因此，d统计量无法用来判定那些通过原点的回归模型的自相关问题。[1]

(2) 变量X是非随机变量，也就是说，在重复取样中是固定的。

(3) 扰动项u_t的产生机制是：

$$u_t=\rho u_{t-1}+v_t \qquad -1\leq\rho\leq 1 \tag{12-7}$$

这表明在t时期内的扰动项或者说误差项依赖于它的第$(t-1)$期的值以及一个纯粹的随机项(v_t)，依赖过去值的程度是由ρ来测度。ρ称为自相关系数(coefficient of autocorrelation)，它介于-1和1之间。(注意：自相关系数总是处于-1和1之间的。)方程(12-7)中所描述的机制，称为马尔可夫一阶自回归(Markov fist-order autoregressive scheme)或者简称为一阶自回归，通常标记为AR(1)方案[AR(1)scheme]。自回归这一名称是恰当的，因为方程(12-7)可以解释为u_t对其滞后一期值的回归。这里是一阶，因为只涉及u_t和它的上一期值，也就是说，最大间隔是一个时期。[2]

(4) 在回归方程中，并没有把应变量的滞后值作为解释变量。换言之，该检验对下面的模型是不适用的：

$$Y_t=B_1+B_2X_t+B_3Y_{t-1}+u_t \tag{12-8}$$

其中Y_{t-1}是应变量Y的一期滞后值。形如式(12-8)的回归方程称为自回归模型(autoregressive models)——变量以其一期滞后值作为解释变量而进行的回归。(我们将在第14章简单讨论这类模型)

假设以上条件都满足，则根据进口支出回归方程的d值0.595 1，能否对自相关作出判断呢？在回答这个问题之前，我们可以证明对大样本来说，式(12-5)可以近似地表达为(见题12.20)：

$$d\approx 2(1-\hat{\rho}) \tag{12-9}$$

并且：

$$\hat{\rho}=\frac{\sum_{t=2}^{n}e_te_{t-1}}{\sum_{t=1}^{n}e_t^2} \tag{12-10}$$

即为自相关系数ρ的估计量。由于$-1\leq\rho\leq 1$，因此，式(12-9)包含如下关系：

ρ值	d值(近似)
1.$\rho=-1$(完全负相关)	$d=4$
2.$\rho=0$(无自相关)	$d=2$
3.$\rho=1$(完全正相关)	$d=0$

简言之，

$$0\leq d\leq 4 \tag{12-11}$$

也就是说，计算的d值必然介于0与4之间。

根据以上讨论我们可以得出：如果计算的d值接近于零，则表明存在着正的自相关，如果接近于4，则表明存在着负的自相关。d值越接近于2，则说明越倾向于无自相关。当然，这些

1 然而，R. W. Farebrother已计算出模型中不存在截距项时的d值。参见，“The Durbin-Watson Test for Serial Correlation When There Is No Intercept in the Regression,”*Econometria*,vol.48，1980，p.1553－1563.

2 如果说模型是

$$u_t=\rho_1u_{t-1}+\rho_2u_{t-2}+v_t$$

则它将是AR(2)或者说二阶自回归过程，如此等等。注意：除非我们假设u的生成过程，否则很难解决自相关问题。这类似于异方差的情形，在异方差情况下，我们对不可观察的误差方差σ_i^2的生成也作了一些假设。对于自相关，在实践中，AR(1)假设证明是非常有用的。

只是很宽泛的临界点，当我们把计算的 d值看作是能够明确说明是正的，负的或无自相关的指标时，我们需要更为具体的临界值。换言之，是否存在一个与 t分布和F分布中那样的临界值，从而对自相关作明确的判定？

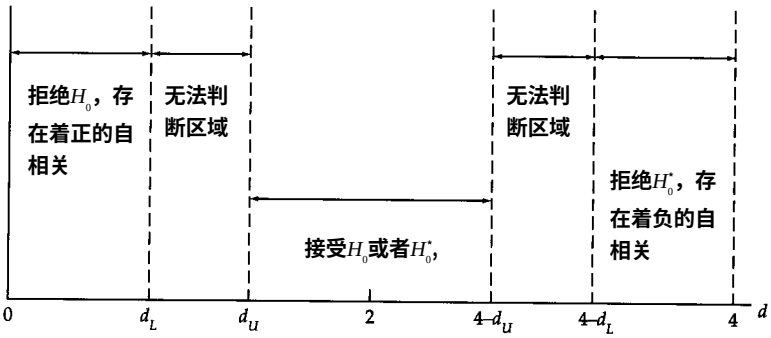


图12-5 杜宾－瓦尔森d统计量

遗憾的是，与t分布或F分布不同，这里有两个而不是一个临界的d值。[1]杜宾和瓦尔森给出了下限d_L和上限d_U，所以，如果根据方程(12-5)计算出的d值位于这些界限之外，我们便可以断定是否存在正的或负的序列相关。这些上限和下限，或者说上临界值和下临界值取决于观察值的个数n，和解释变量的个数k。n可取到从6到200；k最大可达20，杜宾和瓦尔森已经给出了在1%和5%的显著水平下的D-W表，参见附录A，表A-5。我们可用图12-5很好地说明杜宾－瓦尔森检验。

D-W检验的步骤如下：

(1) 进行OLS回归并获得残差e_t。

(2) 根据方程(12-5)计算d值。大多数计算软件已能够实现。

(3) 给定样本容量及解释变量的个数，从 D-W表中查到临界的d_L和e_U。

(4) 按照表12-2中的规则进行判定，为了便以参考，已在图 12-5中对照画出。

表12-2 D-Wd检验：判定规则

零 假 设	决 策	条 件
无正的自相关	拒绝	$0<d<d_L$
无正的自相关	无法决定	$0\leq d\leq d_U$
无负的自相关	拒绝	$4-d_L<d<4$
无负的自相关	无法决定	$4-d_U\leq d\leq 4-d_L$
无正的或者负的自相关	不拒绝	$d_U<d<4-d_U$

回到例12.1中，d=0.595 1。从D-W表中可以看到，对于n=20，k=1，在5%的显著水平下：d_L=1.201和d_U=1.411。由于0.595 1远低于下临界值1.201，根据表12-2中的判定规则，我们可以得出结论：在进口支出回归方程的残差中存在着正的自相关。根据游程检验和图形检验 (图12-3及12-4)我们可以得出同样的结论。

例12.3 抵押债务

为了进一步说明杜宾－瓦尔森 d统计量，我们再来看抵押债务回归方程 (7-32)。根据附录7A.4中所给的计算机输出结果，知道 d=0.402 0。现在对n=16，k=2，临界的d值为0.982和1.539(5%的显著水平下)或0.737和1.252(1%的显著水平下)。由于计算的d值低于这两个下限，因此可以得出结论：残差存在着正的 (一阶)序列相关，从而再一次验证了根据游程检验所得出的结论。

1 我们并不想介绍技术细节，然而需要记住的是， d的关键值取决于解释变量所取的值，而后者显然在样本与样本之间是不同的。

例12.4 对鸡肉的需求

参考方程(10-16)所描述的对鸡肉的需求函数。计算的d值为1.875 6。对于n=23，k=2，临界的d值为1.168和1.543(5%的显著水平下)或0.938和1.291(1%的显著水平下)。因此，我们不能拒绝零假设：数据中不存在自相关。在这个例子中，正的残差数为7，负的残差数为16，游程数为9，根据游程检验也不能拒绝零假设(证明留给读者)。

尽管d检验运用的十分广泛，然而它的一个缺陷是：如果计算得到的d值落入非决策区域或者说是盲区(见图12-5)，那么我们就无法作出是否存在自相关的结论。为了解决这一问题，一些学者[1]提出了对d检验的修正方案，但是相当复杂，已超出了本书的范围。计算机软件SHAZAM可以进行精确的d检验(也就是说真实的临界值)，如果d统计量落入了非决策区域的话，则可以使用这个软件。正如我们看到的那样，自相关的后果可能非常严重，因此，如果d统计量落入了非决策区域，那么假定存在自相关并对条件进行校正可能是鲁莽的。当然，在这种情形下也可以使用非参数检验和图形检验。

在结束d检验的讨论之前，再次需要强调的是：如果d检验本身所需条件都不满足的话，那么就无法使用这种检验方法了。特别需要指出的是：d检验不能对形如式(12-8)的自回归模型进行序列相关检验。如果在这类模型中应用了d检验，则所计算得到的d值通常会在2左右。因此，在这类模型中，检验序列相关就具有内在的偏误。但是如果在实证分析中使用了这样的模型，为了检验自相关问题，杜宾建立了h统计量(h statistic)，详细的讨论参见习题12.16。

12.4 补救措施

由于序列相关可能导致非常严重的后果,并且进一步检验的成本也很高,因此,如果根据前面讨论的一种或者多种诊断检验发现存在自相关问题,则有必要寻找一些补救措施。补救措施取决于我们对误差项u_t相互依赖的性质的了解。为了使讨论尽可能地简单,我们仍以双变量模型为例：

$$Y_t=B_1+B_2X_t+u_t \tag{12-12}$$

假设误差项服从AR(1)过程：

$$u_t=\rho u_{t-1}+v_t \qquad -1\leq\rho\leq1$$

其中，v满足OLS 假定，并且ρ是已知的。

现在如果我们能够对模型(12-12)作变换，使得变换后模型的误差项是序列独立的，然后再用OLS法进行估计，则得到最优线性无偏估计量，当然 CLRM的其他假定也需要满足。在处理异方差问题时我们使用了同样的方法，我们的目标是对模型加以变换，使得在变换后模型的误差项是同方差的。

为了弄清楚如何使变换后模型的误差项不具有自相关性，我们将回归方程 (12-12)中的变量滞后一期，写为：

$$Y_{t-1}=B_1+B_2X_{t-1}+u_{t-1} \tag{12-13}$$[2]

方程(12-13)的两边同时乘以ρ，得到：

$$\rho Y_{t-1}=\rho B_1+\rho B_2X_{t-1}+\rho u_{t-1} \tag{12-14}$$

现在将方程(12-12)与方程(12-14)相减，得到：

1 有些作者认为，杜宾－瓦尔森d统计量的上限d_U近似真实的显著极限。因此，如果计算出的d位于d_U之下，则可以假设存在着(正的)自相关。参见，E. J. Hannan，R. D. Terrell，“Testing for Serial Correlation after Least Squares Regression,”*Econometria*，vol.36，no.2，1968,p.133-150。

2 如果t时期回归方程(12-12)是正确的，则第(t－1)期也必然是正确的。

$$(Y_t - \rho Y_{t-1}) = B_1(1-\rho) + B_2(X_t - \rho X_{t-1}) + v_t \tag{12-15}$$

其中利用了方程(12-7)。

由于方程(12-15)中的误差项v_t满足标准OLS假定，方程(12-15)就是一种变换形式，使得变换后的模型无序列相关。如果我们将方程(12-15)写成：

$$Y_t^* = B_1^* + B_2^* X_t^* + v_t \tag{12-16}$$

其中，

$$Y_t^* = (Y_t - \rho Y_{t-1})$$
$$X_t^* = (X_t - \rho X_{t-1})$$
$$B_1^* = B_1(1-\rho)$$

对变换后的模型使用OLS法，因而获得的估计量具有BLUE性质。同时需要指出的是：对变换后的模型使用OLS得到的估计量称为广义最小二乘(generalized least squares)估计量(GLS)：上一章处理异方差问题时，我们也应用了GLS，只不过在那里，我们称为WLS(加权最小二乘法)。

我们将方程(12-15)或(12-16)称为是广义差分方程(generalized difference equation)。[广义差分方程的特殊情况(即p取特殊值)我们稍后将会讨论到。]广义差分方程包括Y对X的回归，不是用原来的形式，而是以差分的形式。差分形式是通过将变量的当期值减去前期值的一个比例(=p)而得到的。例如，若p=0.5，我们就用该变量当期值减去前期值的0.5倍。在这个差分过程中，由于第一个观察值没有前期值而丢失一个观察值，为了避免丢失这个观察值，可以对Y和X的第一个观察值作如下变换：

$$Y_1^* = \sqrt{1-\rho^2}\,(Y_1)$$
$$X_1^* = \sqrt{1-\rho^2}\,(X_1) \tag{12-17}$$

这一变换称为Prais-Winsten变换(Prais-Winsten transformation)。但是，在实践中，如果样本容量很大，则无需进行这个变换，可以用方程(12-15)(此时有$n-1$个观察值)。

对于广义差分方程(12-15)有几点需要指出：首先，这里我们考虑的仅仅是双变量模型，但是这种差分变换可以推广到多个解释变量的情形(参见习题12.19)。其次，到目前为止，我们仅仅假设了AR(1)，如方程(12-7)，但是差分变换可以很容易地推广到更高阶，例如AR(2)，AR(3)，等等；在变换过程中并不涉及新的问题，只不过计算复杂了一些。[1]

看起来广义差分方程(12-15)中的自相关问题已经有了“解决方法”。不过我们还有一个问题：要成功的应用差分过程，必须知道真实的自相关参数ρ。当然，它是未知的。为了利用方程(12-15)，我们必须通过一些方法来估计未知的ρ。这一情况与异方差情形类似。在那里，我们并不知道真实的σ_i^2，因此不得不做一些可行的假设。当然，如果它是已知的，就可以直接使用加权最小二乘法(WLS)。

12.5 如何估计ρ

ρ的估计方法不是惟一的，下面我们介绍其中的几种。

12.5.1 ρ=1：一阶差分法

既然ρ介于0和±1之间，所以在广义差分方程(12-15)中假设ρ可以取－1到+1之间的任何值。Hildreth和Lu[2]就提出过这样一种方案。但是，究竟是用哪一个特殊的ρ值呢？因为即使是在－1

1 有关高阶自回归过程的讨论可参见，C. Chatfield，*The Analysis of Time Series :An Introduction*，3d ed.,Chapman&Hall，New York，1984.

2 G. Hildreth and J. Y. Lu，“Demand Relations with Autocorrelated Disturbances，”Michigan State University，*Agricultural Experiment Station*，Technical Bulletin 276，November 1960.

和＋1之间，ρ也有成百上千个值可以选择。在应用经济计量学中，广泛采用 ρ=1；也就是说，误差项之间是完全正自相关的，这对一些经济时间序列来说可能是正确的。如果可以接受这个假设，则广义差分方程(12-15)就变为一阶差分方程：

$$Y_t - Y_{t-1} = B_2(X_t - X_{t-1}) + v_t$$

或，

$$\Delta Y_t = B_2 \Delta X_t + v_t \tag{12-18}$$

其中，Δ是一阶差分算子(就像期望算子一样)。在估计方程(12-18)时，首先需要对被解释变量和解释变量同时求差分，然后再对变换后的模型进行回归。

注意一阶差分方程(12-18)的一个重要特征：模型中没有截距。因此，为了估计方程 (12-18)，需要用通过原点的模型(参见题6.15)。很自然，在这种情形下，我们无法直接估计出截距项(但是注意：$b_1 = \bar{Y} - b_2\bar{X}$)。

我们通过进口支出回归方程来说明一阶差分变换。回归的结果与通过其他变换得到的回归结果一起在表12-4中给出，以便比较。稍后我们将对这些变换的结果进行讨论。需要牢记的是：一阶差分变换是明确地建立在 ρ=1这一假定之上的。如果情况不是这样，则不提倡使用这一变换。但是，如何知道ρ的确等于1呢？答案如下。

12.5.2 从杜宾－瓦尔森*d*统计量中估计ρ

前面已建立了*d*与ρ之间的近似关系：

$$d = 2(1 - \hat{\rho}) \tag{12-19}$$

则有：

$$\hat{\rho} = 1 - \frac{d}{2}$$

既然*d*统计量可由大多数回归软件包计算出来，那么，我们可以根据方程 (12-19)很容易地得到ρ的近似估计值。

一旦从方程(12-19)估计出ρ值，我们就可以将其应用于广义差分方程 (12-15)；对进口支出一例，*d*=0.595 1。因此，

$$\begin{aligned}\hat{\rho} &\approx 1 - \frac{0.5951}{2} \\ &\approx 0.7025\end{aligned} \tag{12-20}$$

这一ρ值显然不等于一阶差分方程的假定ρ=1。根据这个ρ值得到的回归结果也在表12-4中给出。

虽说这种方法很容易使用，但只有当样本容量很大时才能得到较理想的 ρ值。对于小样本，Theil和Nagar提出了ρ的另一种估计方法。参见习题 12.21。

12.5.3 从OLS残差e_t中估计*p*

回顾一阶自回归：

$$u_t = \rho u_{t-1} + v_t$$

由于*u*是无法直接观察得到，我们可以用相对应的样本误差 *e*代替，并进行如下回归：

$$e_t = \hat{\rho} e_{t-1} + v_t \tag{12-21}$$

其中 $\hat{\rho}$ 是ρ的估计量。统计理论表明,尽管对小样本而言,$\hat{\rho}$ 是真实ρ的有偏估计量,但是随着样容量的增加,这个偏差会逐渐消失。[1]因此,如果样本容量足够大,可以利用从方程 (12-21)中得到的 $\hat{\rho}$,并用它对数据进行变换,如方程(12-15)所示。方程(12-21)的一个优点在于它简单易行,因为我们利用常用的OLS法就可以获得残差。回归所需的数据在表12-1中给出，回归(12-21)结果如下：

$$\hat{e}_t = 0.739\ 9 e_{t-1}$$

1 技术上，认为 $\hat{\rho}$ 是ρ的一致估计量。

se=(0.193 27)

t=(3.794 7) r^2=0.443 5 (12-22)

因此，所估计得的ρ为0.733 9，与从杜宾－瓦尔森d统计量中所获得的ρ值略有差异。顺便提及，Cochrane-Orcutt法又完善了上述方法。

12.5.4 ρ的其他估计方法

除了我们此前所讨论过的方法以外，还有其他一些估计 ρ的方法：

(1) Cochrane-Orcutt迭代法；(2) Cochrane-Orcut两步法；(3) Durbin两步法；(4) Hidreth-Lu搜索法；(5) 最大似然法。

这里不再对上述方法进行讨论，有兴趣的读者可以查阅有关参考书。[1](见本章的部分习题)无论使用哪种方法，如方程(12-15)所示，我们利用获得的ρ值对数据进行变换并作OLS回归。[2]虽说大多数计算机软件包都能够以最少的指令来完成变换过程，但我们还是在表 12-3中给出变换后的数据；为了说明的方便，我们使用了从方程 (12-22)中获得的ρ值，0.739 9。

在结束本部分的讨论之前，我们来考虑运用如下方法处理进口支出数据所得的结果： (1)运用一阶差分变换， (2)根据d统计量的进行变换， (3)根据方程(12-22)进行的变换。结果在表12-4中给出。关于这些结果有几点需要注意。

(1) 初始回归存在着自相关,但是根据游程检验，各种变换以后的回归不再有自相关问题。[3]

(2) 尽管说，从一阶差分变换中所估计的 ρ值与从d统计量和方程 (12-22)所估计的ρ值并不相同，但是如果回归分析不包括第一个观察值的话，则所估计的斜率和截距系数的差别不大。然而，截距和斜率值却显著地与从通常的 OLS回归中所得的截距和斜率不同。

表12-3 美国进口支出数据：初始的数据和变换后的数据 (ρ=0.7339)

Y_t①	$Y_{(t-1)}$	ρ$Y_{(t-1)}$	$d_1$②	X_t③	$X_{(t-1)}$	ρ$X_{(t-1)}$	$d_2$④
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
135.7	—	(92.248 8)⑤	43.45	155 1.3	—	(105 4.569 6)⑥	496.73
144.6	135.7	99.522 4	45.08	159 9.8	155 1.3	113 7.723	462.08
150.9	144.6	106.049 6	44.85	166 8.1	159 9.8	117 3.293	494.81
166.2	150.9	110.670 0	55.53	172 8.4	166 8.1	122 3.385	505.02
·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·
367.9	351.1	257.496 7	110.40	2 542.8	2 469.8	1 811.351	731.448 7
412.3	367.9	269.817 9	142.48	2 640.9	2 542.8	1 864.890	776.01
439.0	412.3	302.380 8	136.62	2 686.3	2 640.9	1 936.836	749.46

注：如果回归包括第一个观察值，则做 Praid-Winsten变换；如果它不包括在回归中，则舍去第一观察值。表中所有其他剩余项均可以类似方式填入。

① Y_t —— 美国进口支出(1982年，10亿美元)

② d_1 —— 第1列和第3列之差(至小数点后两位数)

③ X_t —— 美国PDI(1982年，10亿美元)

④ d_2 —— 第5列和第7列之差(至小数点后两位数)

⑤ $\sqrt{1-p^2}(Y_1)$

⑥ $\sqrt{1-p^2}(Y_1)$

1 对于这些方法的讨论，参见Damodar N.Gujarati, *Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，p.431-433。

2 对于大样本，使用各种不同的方法所获得的 ρ的估计值的差异会缩小。

3 我们也可以从变换后的回归方程中获得杜宾－瓦尔森 d统计量。但是经济计量理论表明,从变换后的回归中所计算的d统计量可能并不适合于自相关检验,因为如果我们将它用作这一目的,这将表明初始误差项并不遵循AR(1)。它可能遵循,比方说是AR(2)。而游程检验则并不存在着这样的问题,因为它是非参数检验。

表12-4 根据变换后数据得到的进口支出回归结果

回归变换方法	从中所估计的ρ	截距	斜率	R^2	变换后的回归中还存在自相关吗？
初始回归	ρ=0	−261.09	0.2452	0.9388	是的；d=0.5901
	(假设)	t=(−8.3345)	(16.616)		
第一差分	ρ=1		0.328 6	0.9683	否②
	(假设)	t=①	(6.446 8)		
方程(12-15)③	ρ=0.7025	−405.65	0.310 0	0.9729	否②
	(从杜宾的d)	t=(−5.245)	(9.1868)		
方程(12-15)④	ρ=0.7025	−302.07	0.267 4	0.967 9	否②
	(从杜宾的d)	t=(−5.606)	(10.635)		
方程(12-22)③	ρ=0.7339	−430.03	0.320 2	0.973 7	否②
		t=(−5.096)	(8.788 4)		
方程(12-22)④	ρ=0.7339	−307.37	0.2702	0.9682	否②
		t=(−5.3982)	(10.189)		

注：括号中的数值为t值。

所有回归都是由SHAZAM计算机软件包来完成的。

① 在这一模型中没有截距。(为什么？)

② 根据对于所估计得的残差的游程检验。

③ 不包括第一观察值。

④ 包括第一观察值(也就是Prais-Winsten变换)。

(3) 然而，如果通过Prais-Winsten变换纳入第一个观察值，则情况会有显著的变化。变换后的斜率系数非常接近于初始的OLS的斜率，并且变换后模型的截距也非常接近于初始截距。因此，对于小样本而言，包括第一个观察值是很重要的。否则，变换后模型的估计系数将不会有包括第一个观察值的模型那样有效(也就是说，有较大的标准差)。

(4) 各种回归中的r^2值不能直接比较，因为各模型中的解释变量不是一样的。此外，对于一阶差分模型，通常计算的r^2是没有意义的。

如果我们接受了建立在Prais－Winsten变换之上的回归结果(进口支出一例)，并将这些结果与初始回归结果(存在自相关)相比较，我们会看到，初始的截距和斜率系数的t值(绝对值)在变换后的回归中有所下降。换句话说，初始模型低估了标准差。但是根据自相关的理论后果，这一结果并不令人感到吃惊。幸运的是，在这个例子中，即使校正了自相关，估计的t值也依然是统计显著的。[1]但是情况并不总是这样。

12.6 小结

本章要点如下：

(1) 在存在自相关情形下OLS估计量，尽管是无偏的，但不是有效的。简言之，它们不是最优线性无偏估计量。

(2) 在马尔可夫一阶自回归，AR(1)假定下，通常计算的OLS估计量的方差和标准差可能存在着严重的偏差。

(3) 结果，标准t显著性检验和F显著性检验可能存在着严重的误导性。

(4) 因此，重要的是知道在具体情况下是否存在着自相关问题。在所有诊断自相关的方法中，我们考虑了如下三种：

1 严格地说，如果样本容量很大，这一陈述则是正确的。这是因为我们并不知道真实的ρ值，利用估计的ρ值来变换数据时，经济理论表明，在大样本情形下，通常的统计检验过程是有效的。

a. 残差图形检验

b.游程检验

c.杜宾－瓦尔森*d*检验

(5) 如果发现了自相关，我们建议通过适当的变换消除自相关。我们用几个具体的例子说明了实际机制。

习题

12.1 简要解释下列名词：

(a) 自相关 (b) 一阶自相关 (c) 空间相关

12.2 马尔可夫一阶自相关假定有何重要意义？

12.3 在AR(1)假定下，古典线性回归模型的假定之一，总体概率分布函数中的误差项不相关的后果是什么？

12.4 在存在AR(1)自相关的情形下，什么估计方法能够产生 BLUE估计量？简述这个方法的具体步骤。

12.5 在存在AR(1)的情形下，估计自相关参数ρ有哪些不同的方法？

12.6 诊断自相关有哪些不同的方法？说明每一种方法所隐含的假设。

12.7 杜宾－瓦尔森*d*统计量虽广泛使用，但它有什么缺陷？

12.8 判断正误并说明理由。

(a) 当存在自相关时，OLS估计量是有偏的并且也是无效的。 (b) 在形如方程(12-8)的自回归模型中，也就是模型中的一个解释变量是被解释变量的一期滞后值，此时杜宾－瓦尔森 *d* 统计量检验是无效的。 (c) 杜宾－瓦尔森*d*统计量检验假设误差项的同方差。 (d) 消除自相关的一阶差分变换假定自相关系数ρ必须等于1。 (e)两个模型，一个是一阶差分形式，一个是水平形式，这两个模型的R^2值是不可以直接比较的。

12.9 Prais-Winsten变换有什么重要意义？

12.10 完成下表：

样本数	解释变量数	杜宾－瓦尔森*d*统计量	自相关证据
25	2	0.83	有
30	5	1.24	
50	8	1.98	
60	6	3.72	
200	20	1.61	

12.11 利用游程检验检验以下情形中的自相关 (利用Swed-Eisenhart表)

样本数	＋号数	－号数	游程数	自相关
15	11	7	2	
30	15	15	24	
38	20	18	6	
15	8	7	4	
10	5	5	1	

12.12 对第8章给出的菲利蒲斯曲线回归(8-30)，估计的*d*统计量为0.6394。

(a) 残差中是否存在一阶自相关？如果有，是正的还是负的？ (b) 如果存在自相关，根据*d*统计量估计自相关系数。 (c) 利用估计的ρ值，对表8-6中的数据进行变换，并估计广义差分方程(12-15)(也就是说，对变换后的数据用 OLS法)。(d) 在(c)中，估计的回归存在自相关

吗？你使用的是哪一种检验？

12.13 在研究工人在价值增值中所占份额(也即说劳动力的份额)的变化中，根据1949年到1964年期间美国的数据，[1]得到如下回归结果：(t比例值在括号中)

模型A：$\hat{Y}_t = 0.452\ 9 - 0.0041t$； $r^2 = 0.528\ 4$； $d = 0.825\ 2$

$t =$ $(-3.970\ 8)$

模型B：$\hat{Y}_t = 0.478\ 6 - 0.001\ 27t + 0.000\ 5t^2$； $r^2 = 0.528\ 4$；$d = 1.82$

$t =$ $(-3.272\ 4)$ $(2.777\ 7)$

其中Y=劳动份额，t=时间

(a) 在模型A中存在序列相关吗？模型B呢？ (b) 如果在模型A中存在序列相关而模型B中却并不存在，则前者存在序列相关的原因何在？ (c) 这个例子告诉我们在自相关的检验中，d统计量有哪些优点？

12.14 杜宾两阶段法估计ρ。[2]将广义差分方程(12-15)写成如下形式：

$$Y_t = B_1(1-\rho) + B_2X_t - \rho B_2X_{t-1} + \rho Y_{t-1} + v_t$$

第一阶段，杜宾建议以Y作为被解释变量，X_t，X_{t-1}和Y_{t-1}作为解释变量来进行回归。Y_{t-1}的系数将会提供ρ的一个估计量。因此估计得到的ρ是一致估计量；也就是说，对大样本，它是真实ρ值的一个较好估计值。

第二阶段，利用从第一阶段中获得的ρ对数据变换，并估计广义差分方程(12-15)。

将杜宾两阶段法应用于本章所讨论过的美国进口支出数据，并将得到的结果与表12-4中得的结果加以比较。

12.15 考虑如下回归模型：[3]

$$\hat{Y}_t = -49.466\ 4 + 0.885\ 44X_{2t} + 0.0925\ 3X_{3t}；R^2 = 0.997\ 9；0.875\ 5$$

$$t = (-2.239\ 2)\quad (70.293\ 6)\quad (2.693\ 3)$$

其中，Y=个人消费支出(1982年10亿美元)

X_2=个人可支配收入(1982年10亿美元)(PDI)

X_3=道·琼斯工业平均股票指数

根据1961年到1985年期间美国数据。

(a) 在回归的残差中存在一阶自相关吗？你是如何知道的？

(b) 利用杜宾两阶段过程，将上述回归转换成方程(12-15)，结果如下：

$$Y_t^* = -17.97 + 0.89X_{2t}^* + 0.09X_{3t}^*；R^2 = 0.981\ 6；d = 2.28$$

$$t = (30.72)\quad (2.66)$$

自相关的问题解决了吗？你是如何知道的？

(c) 比较初始回归和变换后的回归，PDI的t值急剧下降，这一变化表明了什么？

(d) 根据变换后模型获得的d值能否确定变换后的数据存在自相关？

12.16 杜宾h统计量。在形如方程(12-8)的自回归模型中：

$$Y_t = B_1 + B_2X_t + B_3Y_{t-1} + v_t$$

通常d的d统计量无法用来诊断自相关。对于这类模型，杜宾建议用h统计量来代替d统计量，h统计量定义为：

---

1 参见Damodar N.Gujarati，“Labour’s Share in Manufacturing Industries,”*Industrial and Labor Relations Review*,vol.23，no.1，October 1969，p.63-75.

2 J.Durbin,“Estimation of Parameters in Time-Series Regression Models,”*Journal of the Royal Statistical Society*，ser.B,vol.22，1960，p.139-153.

3 参见Dominick Salvator，Managerial Economics，McGraw-Hill，New York，1989，p.138,148.

$$h = \hat{\rho}\sqrt{\frac{n}{1 - n \ \mathrm{var}(b_3)}}$$

其中，n —— 样本数

$\hat{\rho}$ —— 自相关系数ρ的估计量

$\mathrm{var}(b_3)$ —— 滞后变量Y的系数B_3的估计量的方差

杜宾已经证明，对于大样本而言，给定$\rho=0$这一零假设，h统计量的分布服从：

$$h=N(0,1)$$

也就是说，它服从标准正态分布，也即以零为均值，以 1为方差的正态分布。因此，如果所计算的h统计量超过了h值的临界值，则拒绝$\rho=0$这一零假设；如果它没有超过临界的 h值，则不能拒绝无(一阶)自相关这一零假设。顺便提及，进入 h公式中的$\hat{\rho}$ 可以利用本章所讨论过的任意一种方法获得。

现在考虑1948～1949年到1964～1965年期间印度的货币需求函数如下：

$$\ln\hat{M}_t = 1.602\ 7 - 0.102\ 41\ln R_t + 0.686\ 9\ \ln Y_t + 0.528\ 4\ \ln M_{t-1}$$

$$\mathrm{se}=(1.240\ 4)\ (0.367\ 8)\quad (0.342\ 7)\quad (0.200\ 7)\quad R^2=0.922\ 7$$

式中，M —— 实际现金余额　R —— 长期利率　Y —— 实际国民收入

(a) 对于这一回归，求h统计量并检验假设：上述回归中并不存在一阶自相关。　(b) 对于同一回归，杜宾－瓦尔森 d统计量为1.862 4。说明为什么在这种情况下并不适合用 d统计量。但注意你可以使用 d统计量估计$\rho(\hat{\rho}=1-d/2)$。

12.17 考虑表12-5中所给数据：

表12-5 美国股票价格指数和GNP数据

年份	Y	X	年份	Y	X
1970	45.72	1 015.5	1979	58.32	2 508.2
1971	54.22	1 102.7	1980	68.10	2 732.0
1972	60.29	1 212.8	1981	74.02	3 052.6
1973	57.42	1 359.3	1982	68.93	3 166.0
1974	43.84	1 472.8	1983	92.63	3 405.7
1975	45.73	1 598.4	1984	92.63	3 772.2
1976	54.46	1 782.8	1985	108.09	4 019.2
1977	53.69	1 990.5	1986	136.00	4 240.3
1978	53.70	2 249.7	1987	161.70	4 526.7

注：Y —— NYSE复合普通股票价格指数，(1965年12月31日=100)

X —— GNP(单位：10亿美元)

数据来源：《总统经济报告，1989》，Y列数据来自第416页表B-94，X列数据来自第308页表B-1。

(a) 估计OLS回归：$Y_t=B_1+B_2X_t+u_t$。(b) 根据d统计量确定在数据中是否存在一阶自相关。(c) 如果存在，用d值来估计自相关系数。(d) 利用估计的ρ对数据变换，用OLS法估计广义差分方程(12-15)：(1)舍去第一个观察值；(2)包括第一个观察值。(e) 重复(b)，但根据形如方程(12-21)的残差中估计ρ值。利用估计的ρ值，估计广义差分方程(12-15)。(f) 利用一阶差分法将模型变换成方程(12-18)的形式，并对变换后的模型进行估计。(g) 比较(d)，(e)和(f)的回归结果。你能得出什么结论？在变换后模型中还存在自相关吗？你是如何知道的？

12.18 利用题12-17中所给的数据，得到如下回归结果：

$$\hat{Y}_t = 88.543 - 0.045\ 4X_t + 0.000\ 013\ 1X_t^2 \quad R^2=0.957\ 7；d=1.67$$

$$t=(8.379\ 2)\ (-5.080\ 8)\quad (8.004\ 5)$$

其中X^2表示GNP的平方。

(a) 在这个回归中存在一阶自相关吗？如果不存在，并且在没有包括 GNP平方项的题12.17中发现了自相关，那么你得出什么结论？这两个回归之间有什么差别？ (b) 在这个回归中为什么GNP项是负的而GNP平方项却是正的？这说明了什么？ (c) 先验地，你预期股票价格指数与GNP之间存在着正的还是负的关系？

12.19 考虑如下模型：

$$Y_t=B_1+B_2X_{2t}+B_3X_{3t}+B_4X_{4t}+u_t$$

假设误差项遵循 AR(1)，也即方程 (12-7)。如何变换模型使得变换后的模型不存在自相关？[提示：扩展方程(12-15)]

12.20 建立方程(12-9)(提示：扩展方程(12-15)并利用方程(12-10)。并且注意到对于大样本而言，$\sum e_{t-1}^2$ 与 $\sum e_t^2$ 大致相同。)

12.21 以d统计量之上的 Theil－Nagar的ρ。Theil－Nagar[1]认为，在小样本中不应用 (1－d/2)来估计ρ，而是：

$$\hat{\rho}=\frac{n^2(1-d/2)+k^2}{n^2-k^2}$$

其中，n——样本数；

d——杜宾－瓦尔森d统计量；

k——要估计的系数个数(包括截距项)。

证明对于较大的n而言，这一ρ的估计值等于以简单公式(1－d/2)所得的ρ值。

12.22 蒙特卡罗(Monte Carlo)试验。考虑如下模型：

$$Y_t=1.0+0.9X_t+u_t \tag{1}$$

其中X取值为1，2，3，4，5，6，7，8，9，10。假设有：

$$\begin{aligned}u_t&=\rho u_{t-1}+v_t\\&=0.9u_{t-1}+v_t\end{aligned} \tag{2}$$

其中v_t～$N(0,1)$，假设$u_0=0$。

(a) 生成10个v_t值，再根据方程(2)生成10个u_t值。(b) 利用10个X值和上一步所生成的10个u_t值生成10个Y值。(c) 用(b)部分生成的10个Y值对10个X值进行回归，求出b_1和b_2。(d) 将你所计算的b_1和b_2分别与其真实值1和0.9相比较，结果如何？(e) 从这个例子中你得出什么结论？

12.23 继续习题12.22。现在假设ρ=0.1。你观察到了什么结果？从题12.23和12.22中你能够得出什么结论？

1 参见H. Theil和A. L. Nagar，“Testing the Independence of Regression Disturbances”，*Journal of the American Statistical Association*，vol.56，1961，p.793-806.

第 13 章

模型选择：标准与检验

在前面几章中我们考虑了单方程线性回归模型，例如，对 widgets需求函数、菲利蒲斯曲线、进口支出函数。在给出这些模型时，我们隐含地假定了所选择的模型是“对现实的真实反映”；也就是说，它正确地反映了所要研究的现象。更专业地说，我们假定所选模型中不存在设定偏差或者设定误差。设定误差的产生是由于估计了“不正确的”模型，尽管是不经意地。然而，在实践中，寻找一个真实正确的模型就好像寻找圣杯一样。我们可能永远都无法知道真实的模型是什么，我们所希望的是寻找一个能够“相对”精确反映现实的模型。

正是由于实践的重要性，我们需要对如何建立经济计量模型进行深入的探讨。特别地，我们考虑下列问题：

(1) “好的”或者“正确的”模型有那些性质？

(2) 假设一个无所不知的经济计量学家已经建立了一个“正确”的模型用以分析某种经济现象。然而，由于数据的可获得性，出于对成本的考虑，或者是疏忽等其他原因，研究人员使用了另一个模型，因此，与“正确”模型相比，就犯了设定误差。那么，在实践中可能会犯哪几种类型的设定误差呢？

(3) 设定误差的后果是什么？

(4) 如何诊断设定误差？

(5) 如果已经犯了设定误差，可以采取哪些补救措施重新回到“正确的”模型？

13.1 “好的”模型具有的特性

离开了参考标准或者指导方针，我们就无法确定在实证分析中所选择的模型是不是“好的”、“恰当的”、或“正确的”。著名经济计量学家哈维(A.C.Harvey) [1] 列出了如下标准。根据这些标准，我们可以判断一个模型的好坏。

节省性(parsimony)。一个模型永远也无法完全把握现实；在任何模型的建立过程中，一定程度的抽象或者简化是不可避免的。简单优于复杂 (Occam’s razor寓意)或者节俭原则表明模型应尽可能的简单。

可识别性(identifiability)。即对给定的一组数据，估计的参数必须具有惟一值，或者说，

1 A.C.哈维(A. C. Harvey)，*The Economic Analysis of Time Series*，Wiley ,New York，1981，pp.5-7。以下的讨论主要依据这一本书。另见D. F. Hendry , J. E. Richard，“On the Formulation of Empirical Models in Dynamic Econometrics，”*Journal of Econometrics*,vol.20，October 1982，pp.3-33.

每个参数只有一个估计值。

拟合优度(goodness of fit)。回归分析的基本思想是用模型中所包括的解释变量来尽可能地解释被解释变量的变化。比如我们可用校正的样本决定系数 R^2来度量拟合度，[1] R^2越高，则认为模型就越好。

理论一致性(theoretical consistency)。无论拟合度有多高，一旦模型中的一个或者多个系数的符号有误，该模型就不能说是一个好的模型。因而，在某种商品的需求函数中，如果价格系数为正(需求曲线的斜率为正！)，或者如果收入系数为负(除非这一商品为劣等商品)，那么回归结果就值得怀疑，即使模型的 R^2值很高，比如说0.92。简言之，在构建模型时，我们必须有一些理论基础来支撑这一模型，“没有理论的测量”经常能够导致非常令人失望的结果。

预测能力(predictive power)。正如诺贝尔奖得主米尔顿·弗里德曼(Milton Friedman)所指出的那样，“对假设(模型)的真实性惟一有效的检验就是将预测值与经验值相比较”。[2] 因而，在货币主义模型和凯恩斯模型两者之间选择时，根据这一标准，我们应该选择理论预测能够被实际经验所验证的模型。

虽然建立“好的”模型没有一个统一的方法，但是我们建议读者在建立经济计量模型时应牢记这些标准。

13.2 设定误差的类型

正如前面所指出的那样，模型应该尽可能简单，它应该包括理论上所建议的关键变量而将一些次要影响因素包括在误差项 u中。本节我们讨论几种导致模型失效的设定误差(specification errors)。

13.2.1 遗漏相关变量：“过低拟合”模型

在本章引言中曾提及，由于种种原因，研究者遗漏了一个或多个本应该包括在模型中的解释变量。那么，这对常用的普通最小二乘(OLS)估计过程会有什么影响呢？

我们来看具体的一例子。在第6章中我们曾给出了从1968年到1987年期间美国进口支出函数，见式(6-52)，在第12章中我们又进一步讨论了这个函数。现在，假定“真实”的进口支出函数如下：

$$Y_t=B_1+B_2X_{2t}+B_3X_{3t}+u_t \tag{13-1}$$

式中 Y——进口支出；

X_2——个人可支配收入(PDI)；

X_3——时间或趋势变量，取值从1，2，…到20。

式(13-1)表明：除了PDI以外，还有一个变量，X_3也影响进口支出。它可能是人口、偏好、技术等因素，我们用一个“包罗万象”的变量——时间或者趋势变量表示这些影响因素。[3]

但是，这里，我们不是估计回归方程(13-1)，而是估计下面的方程：

$$Y_t=A_1+A_2X_{2t}+v_t \tag{13-2}$$

---

1 除了样本决定系数R^2以外，还有其他的标准也可用来判断一个模型的拟合度。对于这些标准的详细讨论，参见G.S Maddala，*Introduction to Econometrics*，Macmillan，New York，1988，pp.425-429.

2 Milton Friedman，“The Methodology of Positive Economics，”in *Essays in Positive Economics*，University of Chicago Press,1953，p.7.

3 在时间序列数据中，一个非常普遍的做法是引入一个趋势变量代表那些无法准确表达的其他变量的影响。但是要注意第8章第4节给出的警告。

式(13-2)与式(13-1)类似，只是去掉了“相关”变量X_3。v与u一样也是一个随机误差项。注意：我们用B_s代表“真实”回归式(13-1)的参数，用A_s表示“不正确设定”回归方程(13-2)的参数：与式(13-1)相比，式(13-2)就错误地设定了模型。因此，如果式(13-1)是正确的模型，那么式(13-2)就犯了从模型中排除重要变量的设定误差。我们将这种设定误差称作遗漏变量偏差(omitted variable bias)。它会产生什么后果呢？我们先用一般的形式陈述遗漏变量的后果，然后再用进口支出数据加以说明。

遗漏变量X_3可能会产生如下后果：

(1) 如果遗漏变量X_3与模型中的变量X_2相关，则a_1和a_2是有偏的；也就是说，其均值或期望值与真实值不一致。[1] 用符号表示，

$$E(a_1) \neq B_1 \qquad E(a_2) \neq B_2$$

其中E表示期望(参见第2章)。事实上，可以证明：[2]

$$E(a_2) = B_2 + B_3 b_{32} \tag{13-3}$$

$$E(a_1) = B_1 + B_3(\bar{X}_3 - b_{32}\bar{X}_2) \tag{13-4}$$

其中，b_{32}是遗漏变量X_3对变量X_2的斜率系数。显然，除非式(13-3)中的最后一项为零，否则，a_2将是有偏估计量，偏差的程度取决于最后一项。如果B_3和b_{32}都是正的，则a_2将会有一个向上的偏差——平均而言，它高估了真实的B_2。另一方面，如果B_3是正的，b_{32}是负的，或是B_3为负，b_{32}为正，则a_2将会有一个向下的偏差——平均而言，它低估了真实的B_2。类似地，如果式(13-4)中的最后一项是正的，则a_1将有一个向上的偏差；如果它是负的，则有一个向下的偏差。

(2) a_1和a_2是不一致的，也就是说，无论样本容量有多大，偏差都不会消失。

(3) 如果X_2与X_3不相关，则b_{32}为零。则根据式(13-3)可以看出a_2是无偏的和一致的。[在第4章中曾指出，如果估计量是无偏的(这是一个小样本性质)，它也是一致的(这是一个大样本性质)。反之，却不一定成立；估计量可以是一致的但却不一定是无偏的。] 但是，a_1仍然是有偏的，除非式(13-4)中的$\bar{X}_3$等于零。即使在这种情况下，下面给出的(4)到(6)结论仍然成立[3]。

(4) 根据式(13-2)得到的误差方差是真实误差方差σ^2的有偏估计量。换言之，从真实模型(13-1)估计得到的误差方差与从错误设定模型(13-2)中估计得到的误差方差不相同；前者是真实σ^2的一个无偏估计量，而后者却不是。

(5) 此外，通常估计的a_2的方差$(=\hat{\sigma}^2 / \sum x_{2i}^2)$是真实估计量$b_2$的方差的有偏估计量。即使是$b_{32}$等于零(也即$X_2$与$X_3$不相关)，这一方差仍然是有偏的，可以证明：[4]

$$E[\text{var}(a_2)] = \text{var}(b_2) + \frac{B_3^2 \sum x_{3i}^2}{(n-2)\sum x_{2i}^2} \tag{13-5}$$

也就是说，a_2方差的期望值并不等于b_2的方差：由于式(13-5)的第二项总为正(为什么？)，因此平均而言，$\text{var}(a_2)$高估了b_2的真实方差。这就意味着它将有一个正的偏差。

(6) 因此，通常的置信区间和假设检验过程也就不再可靠。对式(13-5)而言，置信区间将会变宽，因此，我们可能会“更频繁地”接受零假设：系数的真实值为零(或者其它零假设)。

这里不再给出上述各个结论的证明，我们用进口支出函数一例来说明错误设定模型的后果。

---

1 技术要点：根据“非多重共线性”假设，X_2与X_3是否应该是非相关的？回顾第7章可知，变量之间不存在完全共线性这一假设是仅对总体回归函数(PRF)而言，但却无法保证样本中变量之间不相关。

2 证明参阅Damondar N.Gujarati, *Basic Econometrics*, 3d ed，McGraw-Hill，New York，1995，p.205.

3 然而，先验地，没有理由假设$\bar{X}_3$等于零。

4 证明参见Jan Kmenta，*Elements of Eonometrics*, 2d ed.,Macmillan，New York，1986，pp.444-445。注：只有当$b_{32}=0$时才成立。在我们的这个例子中却不是这样，这一点可以从随后的方程(13-8)中看出来。

例13.1 进口支出函数

利用表6-3给出的数据，再加上时间趋势变量，式(13-1)的回归结果如下：

$$\hat{Y}_t = -859.92 + 0.647\,0X_{2t} - 23.195X_{3t}$$
$$\text{se}=(111.96)\quad(0.0745)\quad(4.2704) \tag{13-6}$$
$$t=(-7.6806)\quad(8.6846)\quad(-5.4316)$$
$$\hat{\sigma}^2=184.05;\quad R^2=0.9776;\quad \bar{R}^2=0.975\,0\quad d=1.36$$

而错误设定式(13-2)的回归结果已由第6章的式(6-52)给出，如下：

$$\hat{Y}_t = -261.09 + 0.2452X_{2t}$$
$$\text{se}=(31.327)\quad(0.0148) \tag{13-7}$$
$$t=(-8.334)\quad(16.5676)$$
$$\hat{\sigma}^2=475.48;\quad R^2=0.9388;\quad \bar{R}^2=0.9354;\quad d=0.5951$$

注意两个回归结果的几个特点：

(1) 错误设定式(13-7)表明：PDI每增加一美元，平均而言，用于进口货物上的支出将会增加24美分；也就是说，进口支出的边际倾向为24美分。而真实模型(13-6)表明：由于考虑到趋势变量的影响，因而，PDI每增加一美元，平均而言，用于进口货物的支出将会增加大约65美分。在这个例子中，错误设定方程低估了真实的边际进口支出倾向，也就是说，它有一个向下的偏差。如果作X_3对X_2的回归，得到斜率系数b_{32}，则很容易得到这一向下的偏差：

$$\hat{X}_{3t} = -25.817 + 0.017\,3X_{2t}$$
$$t=(-23.999)\quad(34.177)\quad r^2=0.9848 \tag{13-8}$$

斜率系数b_{32}为0.017 3。从式(13-6)可知，估计的B_2=0.647 0，B_3=-23.195。因此，根据式(13-3)有[1]：$B_2+B_2b_{32}$=0.647 0+(－23.195)(0.017 3)=0.245 2，与从错误设定式(13-7)中得到的值大致相等。注意：B_3(遗漏变量的真实值)和B_{32}(遗漏变量对模型中变量回归的斜率系数)的乘积决定了偏差的性质，是向上或是向下。因而，错误地从模型中略去变量X_3，如方程(13-2)及方程(13-7)，不仅忽略了X_3对Y的影响(B_3)，而且也忽略了X_3对于X_2的影响(b_{32})。因此，“单独的”变量X_2就不得不肩负起遗漏变量对Y的影响，从而无法表现变量X_2对Y的真实影响(0.647 0对比着0.245 2)。

(2) 截距也是有偏的，但在这里它高估了真实的截距值（注意-261是大于-859的)。

(3) 从两个模型中所估计的误差方差也明显不同，分别是184和475。

(4) 截距和斜率(X_2的)的标准差也明显不同。

所有这些结果都与前面的讨论一致。你会发现，如果根据错误设定方程(13-7)来进行假设检验的话，则得出的结论是令人怀疑的。毫无疑问，从模型中略去相关变量可能产生非常严重的后果。因此，在建立模型时，必须小心谨慎。需要对研究现象中所蕴含的经济理论作深入的了解，从而把相关的变量都包括进模型中。如果模型中未包括这些相关变量，则我们就会“过低拟合”或者“过低设定”模型；换言之，我们遗漏了一些重要的变量。

在进一步讨论之前，注意“正确”模型(13-6)和“不正确”模型(13-7)中的杜宾-瓦尔森d

1 注意技术要点。由于在真实模型中b_2和b_3是真实B_2和B_3的无偏估计量，所以，在模型(13-3)中，用的是参数的估计值；也就是说，事实上我们是用表达式$b_2+b_3b_{32}$而不是(13-3)。由于在真实模型中$E(b_3)=B_3$，所以，这里我们用的是b_3。

值，分别为1.36比0.595 1。为什么会有差异呢？这又表明什么呢？稍后，我们回答这些问题。

例13.2 经纪行业中的规模经济

在例11.8中，我们看到纽约股票交易所(NYSE)认为的规模经济是虚假的而非真实的，原因在于样本中把小公司和大公司混在一起，产生了异方差问题[参见方程(11-32)]：当交易成本除以交易量时，我们通过方程(11-33)表明了所谓的规模经济是如何消失的。

我们可以从设定误差的角度来证明这个例子中所观察到的规模经济是虚假的而非真实的。假设我们继续用回归方程 (11-32)，但现加入了6个虚拟变量代表不同规模的公司。回归结果由表13-1给出。令人吃惊的是，平方项的系数为正 (并且在10%的显著水平下是统计显著的)，表明总成本是以一个递增的速率在增加的，这与 NYSE在回归方程(11-32)中所得的结果形成鲜明的对比。换言之，在经纪行业中存在的是规模不经济。从表13-1中的结果可以看到，几个虚拟变量的系数是统计显著的，这表明由于没有考虑到交易量的影响， NYSE犯了模型设定错误。

13.2.2 包括不相关变量：“过度拟合”模型

有的时候，研究人员会采取“大杂烩”的方式将所有的变量都包括进模型中，也不管它们是不是理论上所需要的。“过度拟合”或者“过度设定”模型 (也就是说包括非必须变量)的逻辑思想是只要包括了理论上的相关变量，那么，包括一个或多个不必要的或“无意义”的变量也不会有太大的影响——非相关变量是指没有具体的理论表明应该把这些变量包括到模型中。这样的非相关变量经常被不经意地包括到模型中，因为研究人员不能确定它们在模型中的作用。如果经济理论不完善，这种现象也会发生。在这种情况下，当然会增加 R^2值(若增加变量系数的t值的绝对值大于1，则校正后的R^2也会增加)，从而增加模型的预测能力。

模型中包括非相关变量会导致什么后果呢？我们仍用简单的双变量和三变量模型加以说明。假设：

$$Y_i=B_1+B_2X_{2i}+u_i \tag{13-9}$$

是正确设定的模型，但是，某研究者却加入了多余的变量 X_3，估计了以下的模型：

$$Y_i=A_1+A_2X_{2i}+A_3X_{3i}+v_i \tag{13-10}$$

这里，设定误差是过度拟合了模型，也就是说，包括进了不必要的变量 X_3，不必要是指先验地，它对于Y没有任何影响。式(13-10)的估计结果如下：

表13-1 NYSE经纪成本函数

解释变量	系 数	t
截距	384.539	1.906①
X_2	26.760	8.120④
X_2^2	0.877×10^{-6}	1.886①
D_1	670.530	1.775①
D_2	1604.478	2.730③
D_3	−202.214	0.185
D_4	5 760.251	2.490②
D_5	2 543.144	0.710
D_6	19 488.135	3.690④
R^2=0.94；n=347个公司		

数据来源：Richard R. West and Seha M. Tinic, “Minimum Commission Rates on New York Stock Exchange Transactions,” *Bell Jouranl of Economics*，Vol.2，no.2，p.593，Autumn 1971.

注：D_1=1，交易量为40 000到100 000。
D_2=1，交易量为100 000到200 000。
D_3=1，交易量为200 000到500 000。
D_4=1，交易量为500 000到1 000 000。
D_5=1，交易量为1 000 000到1 800 000。
D_6=1，交易量超过1 800 000。
X_2=交易数。
① 显著水平为0.10。 ② 显著水平为0.01。
③ 显著水平为0.005。 ④ 显著水平为0.0005。

(1)“不正确”模型(13-10)的OLS估计量是无偏的(也是一致的)。也就是说，$E(a_1)=B_1$，$E(a_2)=B_2$和$E(a_3)=0$(注：由于X_3本不属于真实模型，因而B_2的值预期为零。)

(2) 从回归方程(13-10)中所得的σ^2的估计量是正确估计值。

(3) 标准的置信区间和假设检验仍然是有效的。

(4) 但是，从回归方程(13-10)中估计的a却是无效的——通常，它们的方差比从真实模型(13-9)中估计的b的方差大。因此，建立在a的标准差之上的置信区间比建立在b的标准差之上的置信区间宽，尽管，前者的假设-检验是有效的。但是，估计的系数值没有根据正确模型所估计的真实值那么精确。简言之，OLS估计量是线性无偏估计量，但不是最优线性无偏估计量。

注意：到目前为止，我们讨论了两类不同的设定误差。如果略去某一相关变量(过低拟合的情况)，则模型中剩余变量的系数通常是有偏的和不一致的，估计的误差方差也是不正确的，估计量的标准差是有偏的，因此，通常所用的假设-检验过程是无效的。另一方面，若模型中包括了一个无关变量(过度拟合的情况)，则仍然可以得到无偏的和一致估计量，估计的误差方差也是正确的，标准的假设-检验过程仍然是有效的。模型中包括多余变量的一个主要问题是估计系数的方差会变大，因而，对有关真实参数的推断就没那么精确，因为置信区间扩大了。在某些情况下，我们就会接受零假设：真实的系数值为零，因为置信区间变宽了；也就是说，我们会无法认识到被解释变量与解释变量之间的显著关系。

从上述讨论中得到的一个未获保证的的结论是：包括不相关变量比排除相关变量要好一些。但是我们并不鼓励这样做，因为正如前面所指出的那样，不必要变量的增加会减少估计量的有效性(即更大的标准差)，也可能导致多重共线性问题(为什么？)，更不用说对自由度的损失了。

一般地，最优的方法是仅仅包括那些在理论上对应变量有直接影响的解释变量，并且这些解释变量不能够由模型中其他变量解释。[1]

例13.3 经纪行业中的规模经济

除了方程(11-32)以外，NYSE还估计了下面的回归方程：[2]

$$\hat{Y}_i = 381.229 + 33.344\,5X_i - (2.926\times10^{-6})X_i^2 + (2.687\,3\times10^{-13})X_i^3$$
$$t=(2.090\,0)\quad(18.231\,8)\quad(-1.693\,9)\quad(1.072\,4)\quad R^2=0.934 \qquad (13\text{-}11)$$

其中Y=成本，X=股票交易数。

由于式(13-11)中的立方项是统计不显著的，所以，NYSE认为它是一个多余变量。因此，它选择了回归方程(11-32)来验证其发现：经纪行业存在着规模经济。

1 Michael D. Intriligator，*Econometric Models , Techniques and Applications*，Prentice-Hall，Englewood Cliffs，N.J., 1978，p.189.另外别忘了奥卡姆的剃刀原则。

2 这些结果摘自：Richard R. West与Seha M. Tinic，“Minimum Commission Rates on New York Stock Exchanged，”*Bell Journal of Economics*，vol.2，no.2，Autumn 1971，p.591.

(续)

假设接受NYSE回归结果的表面值，留意包括“不必要的”立方项是如何增加平方项的标准差，从而降低了t值的。这与我们刚刚讨论过的理论后果完全一致。当然，我们应该有所选择地接受NYSE的回归结果，因为我们已经看到，这个回归方程可能存在着异方差问题，或是遗漏变量问题，或者两个问题都存在。因此，计算出来的标准差和t值可能是不可靠的。

13.2.3 不正确的函数形式

我们现在考虑另外一种设定误差，我们用进口支出函数 (13-1)来说明。假设模型所包括的变量Y，X_2，X_3都是理论上正确的变量。现在考虑如下形式的进口支出函数：

$$\ln Y_t = A_1 + A_2 \ln X_{2t} + A_3 X_{3t} + v_t \tag{13-12}$$

在方程(13-1)中的变量也进入回归方程(13-12)中，只是变量间的函数关系有所不同；在回归方程(13-12)中，Y的(自然)对数是X_2(自然)对数的线性函数，也是X_3的线性函数。注意：在方程(13-12)中，A_2度量的是进口支出对PDI的弹性，而在方程(13-1)中，B_2度量的仅仅是进口支出与PDI的变化率(斜率)。显然，这两者是不同的。在第 8章中曾介绍过，要获得 Y对X_2的弹性，必须将B_2乘以X_2/Y，显然，这个弹性依赖于所选择的 X_2与Y的值。然而，对模型(13-12)而言，这个弹性系数是不变的，无论 X_2取何值。而且，在方程 (13-1)中，B_3给出了每个时期内进口支出的(绝对)变化率，而在回归方程 (13-12)中，趋势变量的系数给出了 Y的相对变化率(回顾第8章的半对数模型)。同样，这两者也不是一样的。

现在的一个难题是：如何在模型 (13-1)和(13-12)中进行选择，因为经济理论并没有告诉我们应变量与解释变量之间的函数形式。因此，如果回归方程 (13-12)是真实的模型，而我们却用方程(13-1)来拟合数据，则很可能导致模型设定误差，如果情况相反，同样我们可能会错误设定模型，虽然在这两种情形下都包括了相关的变量。由于没有一个很好的理论基础，因此，如果我们选择了错误的函数形式，则所估计的系数可能是真实系数的有偏估计。

例13.4 对数线性进口需求函数

为了扼要地说明这一点，我们给出根据模型 (13-12)拟合的结果：

$$\begin{aligned} \widehat{\ln Y_t} &= -23.727 + 3.897\,5 \ln X_{2t} - 0.052\,6 X_{3t} \\ \text{se} &= (4.431\,4) \quad (0.603\,1) \quad (0.016\,7) \\ t &= (-5.3542) \quad (6.4623) \quad (-3.154\,0) \end{aligned} \tag{13-13}$$

$$R^2 = 0.976\,3; \qquad \bar{R}^2 = 0.973\,5 \qquad d = 1.29$$

回归结果表明：进口支出对 PDI的弹性约为3.9%，也就是说，在其他条件不变的情况下，个人可支配收入增加 1%，平均而言，进口支出将增加 3.9%。然而，对于线性模型(13-1)而言，计算的Y对X_2的弹性约为5.36，这显然比对数线性模型所得的弹性大得多。[1] 而且，模型(13-6)[模型(13-1)的估计量]表明：进口支出在研究时期内以230亿/年的速率减少，而方程 (13-13)却表明它是以5.3%／年的速率在下降，这两个数值无法比较。并且，方程 (13-6)和(13-13)中的R^2值也无法进行比较，因为这两个模型中的被解释变量是不相同的。

1 这一弹性是根据 $b_2(\bar{X}_2 / \bar{Y})$ 计算而来，其中 $\bar{X}_2$ 和 $\bar{Y}$ 是X_2和Y的样本均值。

这个例子清楚地表明：在建立模型时，不仅把理论上相关的变量包括到模型中很重要，而且这些变量之间的函数形式也是很重要的。在双变量回归中，根据散点图很容易选择被解释变量与解释变量之间的函数形式，然而，一旦我们离开了双变量世界，就不再可能将 Y与所有的解释变量同时在图中画出。毕竟，稿纸只是二维的。在后面有关诊断的讨论中，我们将介绍一种检验工具，博克斯-考克斯(Box-Cox)变换。

尽管还有几种其他类型的设定错误，但是，由于篇幅所限我们不再讨论。上述讨论已足够提醒我们：在建立模型时，必须小心谨慎。当然，也没必要为此而感到绝望，因为经济计量技术主要是以经验为基础，并将随着时间的推移逐渐完善。

13.3 诊断设定误差：设定误差的检验

知道设定误差的后果是一回事，而发现犯这类误差又是另一回事，因为我们并非故意犯这类错误。通常，设定误差是不经意产生的，或是由于理论的薄弱使得无法建立准确的模型，或是由于没有正确的数据来检验理论上正确的模型，或是由于被解释变量与解释变量之间的函数形式从理论上来说就不是很明确。实际的问题并不在于犯了这些错误，而在于如何检测出犯了这类错误。一旦确认已经犯了这类设定偏差，则补救措施也就不言而喻了。例如，如果能够证明模型中遗漏了某个解释变量，显然，有效的补救措施就是将这一变量重新纳入模型，当然这个变量的数据是可以得到的。下面，我们讨论几种检验设定错误的方法。

13.3.1 诊断非相关变量的存在

假设模型包括4个变量：

$$Y_i=B_1+B_2X_{2i}+B_3X_{3i}+B_4X_{4i}+u_i \tag{13-14}$$

如果经济理论表明所有这3个X变量都对Y有影响，那么就应该把它们都包括在模型中，即使实证检验发现一个或多个解释变量的系数是统计不显著的。在这种情况下，不会产生非相关变量问题。但是，有时候仅仅是为了避免产生遗漏变量偏差，我们将控制变量(control variables)引入模型。那么可能会出现这种情况：如果控制变量是统计不显著的，并且从模型中删除这些控制变量并未对点估计值或假设检验的结果有太大的影响，则从模型中去掉这些控制变量可能会使模型更好。但需提醒注意的是：必须经过试验。[1] 假定在模型(13-14)中X_4是这种意义上的控制变量，我们无法确定它是否真的属于模型。一个简单的确认方法是估计回归方程(13-14)并检验b_4(也即B_4的估计量)的显著性。在$B_4=0$这一零假设下，我们知道$t=b_4/\text{se}(b_4)$服从自由度为$(n-4)$的t分布(为什么？)。因此，如果计算的t值没有超过给定显著水平下的t临界值，则不拒绝零假设，在这种情况下，X_4可能就是一个多余变量。[2] 当然，如果拒绝零假设，则该变量就很可能是属于模型的。

例13.5 NYSE回归，方程(13-11)

这就是为什么在NYSE一例中，不用方程(13-11)，而用方程(11-32)的原因。因为在方程(13-11)中，立方项的t值仅为1.072 4：这个回归结果是通过对347个公司拟合得到的。因此，利用正态分布近似t分布(为什么？)，临界的t值为1.96 (在5%的显著水平上)和1.65 (在10%的显著水平上)。(但别忘了，在NYSE回归中存在着异方差问题。)

1 在这种情况下，研究者应该告诉读者原始模型(包括丢掉变量)的结果。

2 我们说“可能”是因为如果解释变量之间存在着共线性，则相关系数的标准差会变大，因此会使估计的t值变小。

但是，假定我们并不确定 X_3和X_4都是相关变量。在这种情况下，我们应该检验零假设：$B_3=B_4=0$。根据第7章中讨论的F检验，这很容易做到。(详细内容参见第14章，14.1节的有约束最小二乘的讨论。)

因此，诊断非相关变量的存在与否并不很困难。重要的一点是，在检验的过程中，我们头脑中要有一个“真正”的模型。有了这个模型，我们就能够通过常用的 t检验和F检验来判定一个或者多个变量是否真正相关。然而需要记住的是：我们不能够用 t检验和F检验来重复地建立一个模型；也就是说，我们不能说最开始 Y是与X_2相关的，因为b_2在统计上是显著的，接着将模型加入X_3，如果b_3是统计显著的，就将这个变量保留在模型中。这样一个过程被称作是逐步回归(stepwise regression)。[1] 我们不推荐用“数据挖掘”(data mining)的方法，因为如果说从一开始X_3就属于模型的话，则它早就应该被包括进去。在初始回归中排除 X_3将会犯相关变量遗漏的错误,并且会带来严重的后果。记住：建模必须以理论为指导，否则会陷入死胡同。

13.3.2 对遗漏变量和不正确函数形式的检验

理论应该成为模型的基础，这就产生了一个问题：什么是理论上正确的模型呢？比如，第8章讨论的菲利浦斯曲线一例，尽管我们预期工资变化率与失业率负相关，但究竟哪种函数关系是正确的呢？

$$Y_t=B_1+B_2X_t+u_t \qquad B_2<0$$

或

$$\ln Y_t=B_1+B_2\ln X_t+u_t \qquad B_2<0$$

或

$$Y_t=B_1+B_2\frac{1}{X^t}+u_t \qquad B_2>0$$

正如本章引言中所说，我们不能明确回答这个问题。实践中我们按照如下步骤去判断：首先根据理论或调查以及先前的工作经验，建立了一个自认为抓住了问题的本质的模型。然后对这个模型进行实证检验，并对回归结果进行仔细的分析。别忘了前面讨论的衡量“好”的模型的标准。在这个阶段，我们才能知道所选模型是否恰当。通常，判定模型是否恰当，主要是根据以下一些参数：

(1) R^2和校正后的R^2($\bar{R}^2$)
(2) 估计的t值
(3) 与预期相比，估计系数的符号
(4) 杜宾-瓦尔森d统计量或者趋势统计量
(5) 预测误差

如果这些结果都很好，则可以接受所选模型，认为它较好地代表了现实情况。

如果结果并不令人满意，或是 R^2值太低，或是只有几个系数是统计显著的，或是符号与预期有误，或是杜宾-瓦尔森d统计量太低(表明在数据中可能存在着正的自相关)，或是预测误差相对较大，那么，我们就要考虑模型是否恰当并寻求补救措施：或许我们省略了某个重要的变量，或是使用了错误的函数形式，或是没有用一阶差分时间系列 (参见第12章)去消除序列相关，等等。为了究其“病因”，我们可以用以下的一些方法。

13.3.3 残差检验

在前面曾指出，残差 e_i(在时间序列中是e_t)的检验是目测自相关或异方差存在的一个好的方

1 有关逐步回归的讨论参见Norman Draper和Harry Smith，*Applied Regression Analysis*, 2d ed., Wiley，1981, Chap.6.

法。但是残差也可以用于模型设定误差的检验，比如检验是否遗漏了某个重要变量或使用了不正确的函数形式。为了理解这一点，我们再来看模型 (13-1)和(13-2)。第一个方程将进口支出与PDI和时间趋势联系起来，第二个方程则去掉了时间趋势变量。如果方程 (13-1)是正确的，也就是说趋势变量X_3确实属于模型，但我们却使用了模型 (13-2)，则隐含地认为模型(13-2)误差项为：

$$v_t=B_3X_{3t}+u_t \tag{13-15}$$

因为它不仅要反映真实的随机项u，而且还要反映变量X_3。难怪在这种情形下根据方程(13-2)所估计的残差会显示出一些系统模式 (就是因为模型(13-12)排除了变量X_3)。这可以从图13-1中清楚地看到，该图描绘了回归方程 (13-2)的残差。残差图清楚地表明了残差并没有显示出随机模式。但是，如果我们作正确模型 [方程(13-1)]的残差图，情况会有所不同 (见图13-2)。

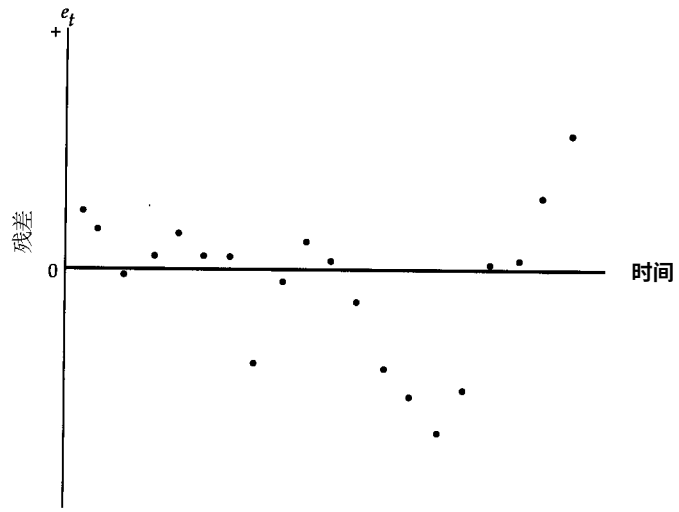


图13-1　回归方程(13-7)的残差

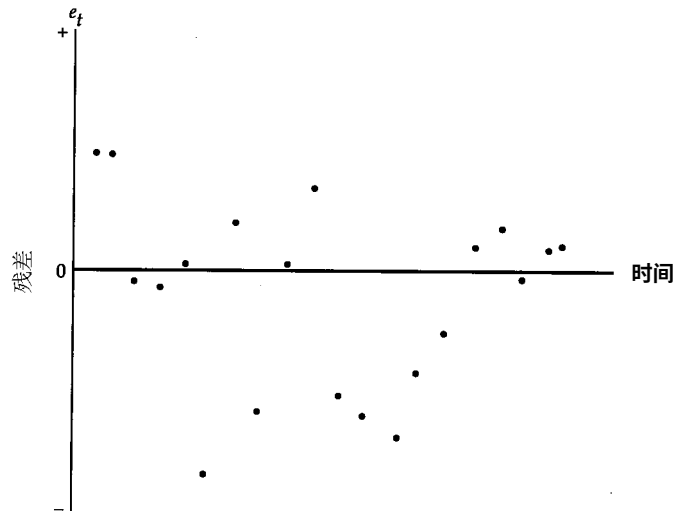


图13-2　回归方程(13-6)的残差

现在残差比以前更随机地分布。但即使是这样，散点图还只是展示了一种可察觉的模式。我们是否省略了其他的变量呢？是否并没有省略某个重要变量，但式 (13-1)的函数形式，或相

对应的式(13-6)，却没有正确设定呢？不幸的是，我们的理论还没有完善到能够告诉我们孰对孰错的地步。

例13.6 进口支出函数的另一种设定形式。假设我们将模型(13-1)做如下修改：

$$Y_t=B_1+B_2X_{2t}+B_3X_{3t}+B_4X_{3t}^2+u_t \tag{13-16}$$

其中，我们加入了趋势变量X_3的平方项。从第8章我们知道，回归方程(13-16)是一个关于X_3的二次多项式。回归结果如下：

$$\hat{Y}_t=-764.68+0.598\,4X_{2t}-26.549X_{3t}+0.291\,2X_{3t}^2$$
$$\text{se}=(89.475)\quad(0.058\,5)\quad(3.391\,0)\quad(0.080\,4) \tag{13-17}$$
$$t=(-8.546\,3)\quad(10.231)\quad(-7.8292)\quad(3.623\,6)$$
$$R^2=0.987\,7;\qquad \bar{R}^2=0.985\,4;\qquad d=2.098\,3$$

对于这个模型的解释是：在其他情况不变的条件下，PDI每增加1美元，平均而言，进口支出增加60美分；也就是说花费在进口货物上的边际倾向为0.60美分，这与从模型(13-6)中所获得的结果在统计上并没有多大区别，后者的结果为0.65美分(你能证明这一点吗？)。在没有趋势平方项的模型(13-6)中，我们看到：平均来说，进口支出是以240亿美元／年这一固定的速率下降的。而在模型(13-7)中，进口支出不是以一个固定的速率下降，而是以一个变化的速率下降。[1]然而有一点值得我们注意，从模型(13-7)得到的残差却是随机分布的，这可以从图13-3中看出来。

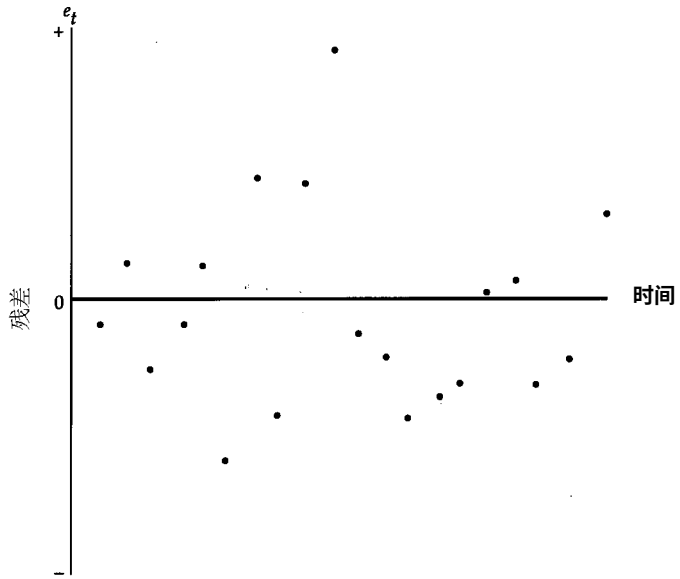


图13-3 从方程(13-7)中所得的残差

1. 再次使用杜宾-瓦尔森d统计量

在第12章中，我们讨论了杜宾-瓦尔森d统计量在检测自相关时的作用。同样的d统计量也可以用于检验设定错误。然而，我们如何能够确定估计的d值是反映了自相关还是反映了设定误差呢？为了回答这一问题，我们来看根据不同形式的进口支出函数所计算得到的d值。

1 如果你有一个像$Y=a+bX+cX^2$这样的一般二次方程，则当X变化时，Y是以一个递增的速率还是以一个递减的速率增加将取决于a，b，c的符号和X的值。关于这些，参见蒋中一，《数理经济学基本方法》，第3版，麦克格罗－希尔出版公司，纽约，1984年，第9章。

模 型	d值	d是否显著？
(13.7)	0.595 1	是①
(13.6)	1.3633	无法确定②
(13.17)	2.098 3	否③

① 对于n=20，k=1，d_L=1.201，d_U=1.411 (5%的显著水平)
② 对于n=20，k=2，d_L=1.110，d_U=1.537 (5%的显著水平)
③ 对于n=20，k=3，d_L=0.998，d_U=1.676 (5%的显著水平)

从表中可以看出，从没有趋势变量的模型到包括趋势变量的模型，再到包括了趋势变量的平方项的模型，在前两个模型中存在自相关问题，而在第三个模型中却消失了。我们可以将这个例子中的d值不仅用于测度自相关，而且还可以测度设定误差。[1]

有一点需要注意，有的时候d统计量，如果显著的话，表明存在着自相关，尤其是当数据是时间序列数据时。但是，它或许表明存在着设定误差。正如我们的例子那样，一个简单的检验方法是：确定原始模型重新设定后，原来显著的d统计量是否会变成不显著的。

2. 其他检验设定误差的方法

残差图和杜宾-瓦尔森d检验都是比较简单的检验方法。还有其他一些复杂程度不一的检验方法，这里，我们只是罗列出来并不详细讨论：[2]

(1) 拉姆齐(Ramsey)RESET检验(设定误差检验)
(2) 似然比例检验
(3) 瓦尔德(Wald)检验
(4) 拉格朗日乘数检验
(5) 霍斯曼(Hausman)检验
(6) 博克斯－考克斯变换(Box-Cox transformation)(以确定回归模型的函数形式)

13.4 用于预测的模型选择

在第1章我们就曾经提到，回归分析的目的之一就是预测应变量的未来值。为了预测，选择模型有那些标准呢？我们用一个具体例子加以说明。习题 11.18中表11-5提供了婴儿死亡率(IMOR)、人均GNP(PCGNP)、初等教育入学年龄组所占百比例 (PEDU)(作为识字率的指标)、人口增长率(POPGROWTH)、以及日常热量摄取量 (CSPC)的数据。假设我们想要根据这些变量中的一个或者多个来预测婴儿死亡率。我们并不确定哪些变量会影响婴儿死亡率。因此，我们考虑下面这些模型：

模型A：IMOR=f(PCGNP，PEDU，POPGROWTH，CSPC)
模型B：IMOR=f(PCGNP，PEDU，CSPC)
模型C：IMOR=f(PCGNP，PEDU，POPGROWTH)
模型D：IMOR=f(PCGNP，PEDU，CSPC)

其中f是函数符号。

上面的哪些模型的预测效果较好呢？也就是说，哪些模型预测的误差最小呢？记住，这里我们不是讨论根据样本内 (in-sample)数据得到的结果。如果是那样的话，我们可以用常用的 R^2

1 参见Damodar N.Gujarati, *Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，p.462。
2 参见Damodar N.Gujarati, *Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，p.462,Chap.13。对于这些检验的较为高级的讨论可以参见： Thomas B Fomby，R Carter Hill，Stanley R.Johnson, *Advanced Econometric Methods*，Springer-Verlag，New York，1984，Chap.18。

值来判断。

有多种判断回归模型预测效果的标准。最常用的两种是 Akaike的信息标准(information criterion)(AIC)和Schwarz的信息标准(SIC)。这些标准的定义如下：

$$\text{AIC}=e^{\frac{2k}{n}}\ \frac{\sum e_t^2}{n} \tag{13-18}$$

$$\text{SIC}=n^{k/n}\frac{\sum e_t^2}{n} \tag{13-19}$$

其中，e是自然对数($e\approx 2.718\ 3$)，n是样本中观察值的总数，k是模型中变量的总数(包括截距项)，$\sum e_t^2$是样本残差平方和(RSS)。这些公式都给出了预测误差方差的估计值。(有关预测误差的细节，参见第6章第6.10节。)

上述两个公式前面的因子可认为是自由度处罚因子(freedom penalty factor)[1]：模型中解释变量的个数越多，则处罚越重。相对来说，SIC标准对自由度的处罚比AIC标准更重(你能验证这一点吗？)无论是用AIC标准还是SIC标准，从预测的角度来看，度量值越低，模型的预测会越好(为什么？)。

再来看上面的4个模型，利用EVIEWS统计软件得到各个模型的AIC值和SIC值。(注意：EVIEWS给出的是AIC和SIC的对数值。但是由于在两种测度是一一对应的关系，因而不会产生太大的问题。)

模型	lnAIC	lnSIC
A	7.331 9	7.580 9
B	7.220 2	7.369 5
C	7.262 7	7.461 8
D	7.258 7	7.457 8

由于模型B中AIC值和SIC最低，所以，从预测的角度来看，模型B最好。

13.5 小结

本章讨论的要点如下：

(1) 古典线性回归模型假定实证分析中所使用的模型是“正确设定的”。

(2) 模型的正确设定有几种含义，例如：

a. 模型没有排除理论上的相关变量。

b.模型没有包括非相关变量。

c.模型的函数形式是正确的。

(3) 如果一个或多个理论上的相关变量被排除在模型之外，则模型中剩余解释变量的系数通常是有偏的和不一致的，OLS估计量的误差方差和标准差也都是有偏的。因此，至少可以说，传统的t检验和F检验的结果是不可靠的。

(4) 使用错误的函数形式，会有类似的后果。

(5) 相对来说，如果模型中包括了非相关变量，则后果没有那么严重，主要表现在：估计的系数仍然是无偏的和一致的，估计量的误差方差和标准差也是正确的，而且，传统的假设检验过程仍然是有效的。但估计的标准差会相对变大，这意味着模型中的参数估计值并不是很精确。从而导致置信区间会变宽。

1 这一术语出自Francis X. Diebold，*Elements of Forecasting*，South Western Publishing Company，Cincinnati，Ohio，1998，p.24。

(6) 本章讨论了几种诊断工具以帮助我们确定在具体应用中是否存在设定错误。这些工具包括残差的图形检验和杜宾-瓦尔森d统计量检验。我们还提到了其他的一些检验方法，如拉姆齐(Ramsy)的RESET检验，拉格朗日乘数检验，霍斯曼(Hausman)检验和博克斯-考克斯(Box-Cox)变换。[1]

(7) 最后，本章还讨论了评估模型预测能力的AIC标准和SIC标准。这些标准都主要针对样本外的预测误差：样本外的预测误差越小，则模型越好。

寻找理论上正确的模型可能相当麻烦，但需牢记几个实践标准，例如：(1)过度节俭；(2)可识别；(3)拟合优度；(4)理论一致性；(5)预测能力。

建模既是一门艺术也是一门科学。仅靠经济计量学知识和计算机统计软件并不足于确保我们成功。[2]

习题

13.1　设定误差的含意是什么？

13.2　设定误差产生的原因是什么？

13.3　“好的”经济计量模型有哪些属性？

13.4　设定误差有哪些类型？它们是否可能同时发生？

13.5　模型遗漏相关变量的后果是什么？

13.6　我们说一个变量是“相关”或者是“不相关”，这意味着什么？

13.7　模型中包括非相关变量的后果是什么？

13.8　“从模型中遗漏一个或者多个相关变量比在模型中包括一个或者多个非相关变量的后果更为严重。”对此，你是否同意，为什么？

13.9　在求简单凯恩斯乘数的过程中，作GNP对投资回归，发现它们之间存在着某种关系。现在，你认为在模型中包括了“非相关”变量“国家和地方税收”，它对模型并无太大影响。但出乎意料的是，投资变量却不显著了。一个非相关变量是如何做到这样的？

13.10　如果你将在以下两个模型中进行选择，那么你将选择哪一个？一个模型满足所有的统计标准但却并不满足经济理论，另一个模型符合已有的经济理论但却不能满足几项统计标准。

13.11　考虑表8-7中给出的假想的成本与产出数据。根据这些数据，我们获得了以下几个成本函数(见表13-2)：

表13-2　根据表8-7中的数据得到的成本函数

成本函数类型	截　距	X	X^2	X^3	R^2	d
线性	166.467	19.933	—	—	0.840 9	0.716
	se=(19.021)	(3.066)				
平方型	222.383	-8.025	2.542		0.928 4	1.038
	se=(23.488)	(9.809)	(0.869)			
立方型	141.767	63.478	-12.962	0.939	0.998 3	2.70
	se=(6.375)	(4.778)	(0.986)	(0.059)		

1　有关这些及其他检验方法的讨论参见L. G. Godfrey，*Misspecification Tests in Econometrics: The Larange Multiplier Principle and Other Approaches*，Cambridge University Press，New York，1988。

2　C. W. J. Granger编著的，*Modeling Economic Series: Readings in Econometric Methodology*, Clarendon, Oxford, U.K.1990,p.2.。这是一本非常有意思的书，它给出了主要的一些经济计量学家们有关建模的艺术与科学的观点。

(a) 观察值为10，解释变量个数分别为1，2和3时，在α=5%的显著水平上，杜宾－瓦尔森 d 统计量的临界值各是多少？ (b) 在线性成本函数中是否存在着“自相关”？如果存在，意味着什么？ (c) 在平方型成本函数中是否存在着“自相关”？如果存在，意味着什么？ (d) 在立方型成本函数中是否存在着自相关，如果不存在，则对在线性和平方型成本函数中所观察到的自相关现象，你作何解释？ (e) 这三个成本函数的边际成本是多少？ (f) 从经济学的角度来衡量，哪个成本函数有意义，为什么？

13.12 表13-3给出了1958年到1972年期间，中国台湾制造业部门实际总成本，劳动投入和资本投入的数据。假设理论上正确的生产函数是科布－道格拉斯函数，如下：

$$\ln Y_t = B_1 + B_2 \ln X_{2t} + B_3 \ln X_{3t} + u_t$$

其中ln为自然对数。

表 13-3

年份	Y	X_2	X_3	X_4
1958	8 911.4	281.5	120 573	1
1959	10 873.2	284.4	122 242	2
1960	11 132.5	289.0	125 263	3
1961	12 086.5	375.8	128 539	4
1962	12 767.5	375.2	131 427	5
1963	16 347.1	402.5	134 267	6
1964	19 542.7	478.0	139 038	7
1965	21 075.9	553.4	146 450	8
1966	23 052.0	616.7	153 714	9
1967	26 128.2	695.7	164 783	10
1968	29 563.7	790.3	176 864	11
1969	33 376.6	816.0	188 146	12
1970	38 354.3	848.4	205 841	13
1971	46 868.3	873.1	221 748	14
1972	54 308	999.2	239 715	15

数据来源：Thomas Pei-Fan Chen，“Economic Growth and Structure Change in Taiwan ,1952-1972,A Production Function Approach，Unpublished Ph.D thesis，Department of Economics，Graduate Center，City University of New York，June 1976，Table 2。

注：Y —— 实际总产出，单位：百万新台币。
X_2 —— 劳动投入，单位：千人。
X_3 —— 实际资本投入，单位：百万新台币。
X_4 —— 时间或者趋势变量。

(a) 估计样本期间内台湾的科布-道格拉斯生产函数并阐述所得的结果。 (b) 假设资本数据一开始无法获得，因此有人估计了以下生产函数。

$$\ln Y_t = C_1 + C_2 \ln X_{3t} + v_t$$

其中 v 为误差项。在这一情形下存在着什么样的设定误差？其后果是什么？用手头的数据加以说明。

(c) 现在假设劳动力投入数据一开始无法获得，假设你估计的是以下模型：

$$\ln Y_t = C_1 + C_2 \ln X_{3t} + w_t$$

其中 w 为误差项。这种设定误差的后果是什么？用表 13-3中的数据来说明。

13.13 继续练习题13.12。由于数据为时间序列，因而在数据中可能存在自相关。用杜宾-瓦尔森 d 统计量来检验这一可能性。如果在数据中的确存在 (一阶)自相关，用估计的 d 值对数据做变换，并根据变换后数据，求回归结果。两个结果是否有质的区别？

13.14　考虑如下模型：

模型I：消费$_i$ $=B_1+B_2$收入$_i+u_i$

模型II：消费$_i$ $=A_1+A_2$财富$_i+v_i$

(a) 你如何确定哪个模型是“正确的”模型？　(b) 假设你同时以收入和财富来对消费进行回归。这如何帮助你确定选择哪个模型？　(c) 你预期在收入与财富之间是否会存在完全共线性？为什么？　(d) 如果在收入与财富之间存在高度但不是完全的共线性，假设把两个变量都用来解释消费，你预期会产生什么后果？在这种情况下，你如何在模型 I和模型II之间取舍？

13.15　参考第6章习题6.15，该题讨论了经过原点的回归(也就是说零截距)模型。如果真实的模型存在截距项，而你却用通过原点的回归模型来估计，这会导致什么样的设定误差？用习题6.15中的数据来说明这一类型误差的后果。

13.16　表13-4给出了1954年到1981年期间美国普通股实际收益率(Y)，产出增长率(X_2)和通货膨胀(X_3)的有关数据，所有数据均以百分比的形式给出。

表　13-4

年份	Y	X_2	X_3	年份	Y	X_2	X_3
1954	53.0	6.7	−0.4	1968	6.8	2.8	4.3
1955	31.2	2.1	0.4	1969	-13.5	-0.2	5.0
1956	3.7	1.8	2.9	1970	-0.4	3.4	4.4
1957	-13.8	-0.4	3.0	1971	10.5	5.7	3.8
1958	41.7	6.0	1.7	1972	15.4	5.8	3.6
1959	10.5	2.1	1.5	1973	-22.6	-0.6	7.9
1960	-1.3	2.6	1.8	1974	-37.3	-1.2	10.8
1961	26.1	5.8	0.8	1975	31.2	5.4	6.0
1962	-10.5	4.0	1.8	1976	19.1	5.5	4.7
1963	21.2	5.3	1.6	1977	-13.1	5.0	5.9
1964	15.5	6.0	1.0	1978	-1.3	2.8	7.9
1965	10.2	6.0	2.3	1979	8.6	-0.3	9.8
1966	-13.3	2.7	3.2	1980	−22.2	2.6	10.2
1967	21.3	4.6	2.7	1981	−12.2	-1.9	7.3

数据来源：Jason Benderly，Burton Zwick，“Inflation , Real Balance, Output and Real Stock Returns”，*American Economic Review*，vol.75，no.5，December 1985，p.1117。

(a) 用X_3对Y进行回归。　(b) 用X_2和X_3对Y进行回归。　(c) Eugene Fama教授指出，“实际股票收益和通货膨胀之间的简单的负相关关系是虚假的(或者说是错误的)，因为它是两个结构关系的结果：一是在当前实际股票收益和预期产出增长率之间的正相关关系，另一是在预期产出增长率和当前通货膨胀之间的负相关关系”。根据这个观点评论以上两个回归结果。(d) 就1956年到1976年期间的数据进行(b)部分的回归，同时省略1954年和1955年的数据，因为这两年的股票收益行为异常。将获得的回归结果与在(b)部分的回归结果相比较，如果两者存在着差异的话，就两者的差异作评论。　(e) 假设你想就1956年到1981年的数据作回归，但是想区分1956年到1976年和1977年到1981年这两个时期，你将如何进行这一回归？(提示：考虑虚拟变量)

13.17　表13-5给出了1960年到1982年期间OECD国家——美国、加拿大、德国、法国、英国、意大利和日本这几个国家的总最终能源需求(Y)，实际国内生产总值，GDP (X_2)和实际能源价格(X_3)等几个变量指数的数据。(所有的指数都是以1973年为100)(另见表8-3)

年份	Y	X_2	X_3	年份	Y	X_2	X_3
1960	54.1	54.1	111.9	1972	97.7	94.3	98.6
1961	55.4	56.4	112.2	1973	100.0	100.0	100.0
1962	58.5	59.4	111.1	1974	97.4	101.4	120.1
1963	61.7	62.1	110.2	1975	93.5	100.5	131.0
1964	63.6	65.9	109.0	1976	99.1	105.3	129.6
1965	66.8	69.5	108.3	1977	100.9	109.9	137.7
1966	70.3	73.2	105.3	1978	103.9	114.4	133.7
1967	73.5	75.7	104.3	1979	106.9	118.3	144.5
1968	78.3	79.9	104.3	1980	101.2	119.6	179.0
1969	83.8	83.8	101.7	1981	98.1	121.1	189.4
1970	88.9	86.2	97.7	1982	95.6	120.6	190.9
1971	91.8	89.9	100.3				

数据来源：Richard D. Prosser，“Demand Elasticities in OECD:Dynamic Aspects,”*Energy Economics*,January1985,p.10.

(a)**估计以下模型：**

模型A：$\ln Y_t=B_1+B_2\ln X_{2t}+B_3\ln X_{3t}+u_{lt}$

模型B：$\ln Y_t=A_1+A_2\ln X_{2t}+A_3\ln X_{3(t-1)}+A_4\ln X_{3t}+u_{2t}$

模型C：$\ln Y_t=C_1+C_2\ln X_{2t}+C_3\ln X_{3t}+C_4\ln X_{(3t-1)}+u_{3t}$

模型D：$\ln Y_t=D_1+D_2\ln X_{2t}+D_3\ln X_{3t}+D_4\ln Y_{(t-1)}+u_{4t}$

其中u**为误差项。注意模型**B**和**C**称为动态模型——明确考虑了不同时期变量的变化。模型**B**和**C**称为分布滞后模型，因为一个解释变量对于被解释变量的影响扩散到了多个时期，这里是两个时期。模型** D**称为自回归模型，因为其中的一个解释变量是被解释变量的一个滞后值。(动态模型将在第**14**章中予以讨论)**

(c) **如果你只估计了模型**A**，而真实模型是**B**、**C**或**D**，则会犯什么样的设定误差？**

(d) **由于以上模型都是线性对数形式的，因此，其斜率系数也就代表了弹性系数。模型** A **的收入弹性(也就是说对于**GDP**而言)和价格弹性是多少？你将如何估计其它三个模型的弹性？**

(e) **由于在模型**D**中滞后的**Y**变量以解释变量的形式出现，在对模型** D**进行**OLS**估计时你预计会出现什么问题？(提示：回顾**CLRM**的假定条件)**

13.18 **证明从方程**(13-6)**中得到的边际进口倾向，**0.647 0**和从方程**(13-17)**中得到的边际进口倾向，**0.598 4**不是显著不同的。**

13.19 **参考习题**13.12**。假设加入代表技术的趋势变量** X_4 **来扩展科布-道格拉斯函数模型。进一步假设** X_4 **是统计显著的。在这种情形下，犯了哪种设定误差？如果** X_4 **在统计上是不显著的，则又会如何？给出必要的计算。**

13.20 **对鸡肉需求的考虑。如果方程** (10-15)**是正确的鸡肉需求函数，而我们却用了方程**(10-15)**，则会导致何种设定误差？这一设定误差的后果是什么？给出必要的解释。**

13.21 **在百兹摩星期五，**478**美国**385**宗案卷**(1986**年**)中，一宗涉及到北卡罗莱纳延长服务公司工资歧视的案卷，原告为一群黑人工人，他们给出了一个多重回归模型以表明：平均而言，黑人工人所获得的薪水低于白人工人所获得的薪水。当案卷到达地区法院时，法庭却拒绝了原告的申诉，理由是在他们的回归中并没有包括影响工资的所有变量。然而高级法院却重新扭转了地区法院的判决。它认为：**[1]

1 **摘自**Michael O. Finkelstain**和**Bruce levin，*Statistics for Lawyers*, Springer-Verlag, New York，1989，p.374。

地区法院错误地认为，申诉者们的回归分析“作为歧视的证明是无法接受的”，因为它们并没有包括影响工资的所有变量。地区法院关于回归分析的证据的看法显然是不正确的。尽管从分析模型中遗漏变量会使分析的检验能力比不遗漏变量要低一些，但是却很难说，由于缺少一些次要的因素，用以分析的主要因素就“一定不能作为歧视的证据”(同前)。通常，未能包括所有变量会影响分析的检验能力，但却不会影响它作为证据的资格。

你认为高级法院的判决是正确的吗？详细地说明你的观点，提醒注意设定错误的理论和实际后果。

13.22 根据AIC和SIC标准选择以下两个模型：在第7章第7.4节中所讨论的加入抵押成本的抵押债务模型和不加入抵押成本的模型。

第 14 章

■ 单方程回归模型：几个补充专题

本章我们介绍在实际研究中非常有用的 4个专题。这些专题是：

(1) 有限最小二乘法

(2) 动态经济模型

(3) 分对数模型

(4) 伪回归

我们将通过具体的几个例子来加以说明。

14.1 有限最小二乘法

在估计线性回归模型的参数时，我们使用了普通最小二乘法 (OLS)。OLS法的基本思想是估计未知参数使残差平方和 (RSS) $\sum e_i^2$ 最小。在这个最小化过程中，并没有对参数加任何限制。例如，在三变量模型

$$Y_i=B_1+B_2X_{2i}+B_3X_{3i}+u_i$$

中，我们并未假设 $B_2=2$(或者其他一些值)或者$(B_2+B_3)=1$(或是其他限制)。换言之，至今为止，我们是“ 自由地 ”估计这些变量的，也即没有任何限制：无论 $\sum e_i^2$ 最小值是多少，我们都予以接受。换句话说，我们通过给定样本获得回归参数的最优值，而不加以任何限制或者约束。因此，迄今为止，我们所使用的最小二乘法可称为无限制条件的最小二乘法 (unrestricted least squares) (ULS)。

但是，有些时候，经济理论表明在回归模型中的系数满足一定的限制条件。例如，前面遇到过的科布－道格拉斯生产函数：

$$\ln Y_i=B_1+B_2\ln X_{2i}+B_3\ln X_{3i}+u_i \tag{14-1}$$

式中 Y—— 产出

X_2—— 劳动投入

X_3—— 资本投入

ln —— 自然对数

现在，如果规模收益不变 (例如，投入增加一倍，产出也同时增加一倍)，根据经济理论有：

$$B_2+B_3=1 \tag{14-2}$$

也就是说，两个产出-投入弹性之和等于 1。[1] 则方程(14-2)就是一个线性恒等约束 (linear equality restriction)的例子。[2]

我们如何知道在具体研究中，科布－道格拉斯型生产函数存在规模收益不变呢？也就是说，如何知道约束方程 (14-2)是有效的呢？更具体地，我们来看习题 13.12给出的台湾制造业的投入-产出数据。在习题 13.12中，要求用科布－道格拉斯生产函数拟合数据。拟合结果如下 (也就是说，无约束最小二乘法)：

$$\ln \hat{Y}_t = -8.401\ 0 + 0.673\ 1 \ln X_{2t} + 1.181\ 6 \ln X_{3t}$$
$$\text{se} = (2.717\ 7) \quad (0.153\ 14) \quad (0.302\ 04) \quad R^2 = 0.9824 \tag{14-3}$$
$$t = (-3.091\ 2) \quad (4.395\ 2) \quad (3.912\ 1)$$

现在，估计的弹性之和为 $(b_2+b_3)=(0.673\ 1+1.181\ 6)=1.854\ 7$。从数学上说，该值大于 1，但是从统计上说，却并不一定大于 1，原因在于抽样的波动性。换言之，线性约束方程 (14-2)可能仍是统计有效的。我们如何确认这一点呢？有几种不同的方法回答这一问题，但是，常用的方法是有限最小二乘法(restricted least squares，RLS)。

若约束方程(14-2)的确是有效的，则有：

$$B_2 = 1 - B_3 \tag{14-4}$$

(也可以表达为 $B_3=1-B_2$)。因此，我们可将回归方程(14-1)表示成

$$\ln Y_t = B_1 + (1-B_3)\ln X_{2t} + B_3 \ln X_{3t} + u_t = B_1 + \ln X_{2t} + B_3(\ln X_{3t} - \ln X_{2t}) + u_t$$

也即，

$$(\ln Y_t - \ln X_{2t}) = B_1 + B_3(\ln X_{3t} - \ln X_{2t}) + u_t \tag{14-5}$$

或

$$\ln \frac{Y_t}{X_{2t}} = B_1 + B_3 \ln \frac{X_{3t}}{X_{2t}} + u_t \tag{14-6}$$

其中，(Y_t/X_{2t})表示产出与劳动的比值(也就是劳动生产率)，(X_{3t}/X_{2t})表示资本与劳动的比值(即每单位劳动力所拥有的资本量)，这些比值具有重要的经济意义。

方程(14-6)表明，劳动生产率的对数是资本／劳动比对数的线性函数。注意方程 (14-6)的几个特点：

(1) 它表明了如何将经济理论所蕴含的限制条件，例如方程 (14-2)或方程(14-4)直接加入到方程中。

(2) 它只要求估计两个未知系数，而方程 (14-1)却要求估计三个未知系数。

(3) 一旦从方程(14-6)中估计出 B_3，则可以很容易从方程 (14-4)中估计出 B_2。显然，这个过程能够保证 $B_2+B_3=1$。(为什么？)

方程(14-5)或者(14-6)所概述的过程称为有限最小二乘法 (RLS)，因为我们根据相关理论考虑到了某些限制条件，并将数据加以变换后才应用最小二乘法。前面已经指出，通常的 OLS法可以称作无约束最小二乘法 (ULS)。

利用回归方程(14-1)和(14-6)如何检验约束方程(14-4)的有效性呢？为了回答这个问题，我们首先给出式(14-6)的回归结果。

1　B_2是劳动投入对产出的弹性，它度量的是在资本投入保持不变的条件下，劳动投入每变化百分之一，产出变化的百分比。我们知道在双对数模型中，斜率系数也是弹性。

2　如果 $B_2+B_3<1$，就是线性不等约束 (linear inequality restriction)的例子。解决这种约束问题，需要利用数学规划技术，它已超出本书讨论的范围。

例14.1 台湾的RLS科布-道格拉斯生产函数

$$\ln \frac{\hat{Y}_t}{X_{2t}} = 5.506\ 7 - 0.321\ 56 \ln \frac{X_{3t}}{X_{2t}} \qquad R^{*2} = 0.942\ 8$$

$$\text{se} = (0.880\ 82) \quad (0.154\ 76) \tag{14-7}$$

$$t = (6.251\ 8) \quad (-2.077\ 7)$$

方程(14-7)是一个有限最小二乘回归，因为它考虑了方程 (14-2)所附加的限制条件。注意RLS回归(14-7)中的这些特点：

(1) 由于估计的 B_3为－0.321 5，根据模型 (14-7)很容易得到 B_2的估计值为 1－(-0.321 5)=1.321 5；也就是说估计的产出－劳动弹性为 1.32%。因为，估计的弹性之和为1。(为什么？)

(2) 虽然在回归(14-7)中的b_3是统计显著的(验证一下)，但其值却是负的。它没有什么经济意义，因为经济理论表明，一般而言，资本/劳动比例越高，劳动生产率也越高。

(3) RLS回归中的R^2 (记为R^{*2})值低于原始回归 (14-3)(ULS)中的R^2值。一般来说这是正确的。(参见本例末尾的注释)

与预期相背，估计的 B_3值为负，这表明：对于台湾制造业而言， $B_2+B_3=1$这一约束——它表明规模收益不变——看来是无效的。但这只是一个定性推断。我们可以给出判定约束条件有效性的一个更为直接的检验方法：

令

R^2=无限制回归(14-3)的R^2

R^{*2}=限制回归(14-7)的R^2

m=线性限制数(在此例中为1) [1]

k=无限制回归(14-3)参数的个数(在此例子中为3)

n=观察值的个数

假定误差项u服从正态分布，可以证明：

$$F = \frac{(R^2 - R^{*2})/m}{(1 - R^2)/(n - k)} \sim F_{m,\ (n-k)} \tag{14-8}$$

服从分子和分母自由度分别为m和$(n-k)$的F分布。因此，在具体的应用中，如果从方程(14-8)中计算得到的F值大于给定显著水平下的临界的 F值，则根据理论所附加的特定约束条件是无效的(从统计上说)。另一方面，如果计算得到的F值小于临界的F值，则可以接受：约束条件是有效的。在这种情形下， RLS回归就优于ULS回归。

上述F检验的步骤如下：

(1) 估计ULS回归(也即普通最小二乘(OLS)回归结果)并获得R^2。

(2) 估计RLS回归并获得相应的决定系数，记为 R^{*2}。

(3) 求出约束条件个数 (m)和无约束回归中所要估计系数的个数 (k)，并求出F值。零假设为：约束条件是有效的。

(4) 如果从方程(14-8)所获得的F值没有超过选定显著水平下(也就是说，犯第一类错误的概率)的临界的F值，则可以接受零假设 (也即给定的约束条件是有效的)。但是如果F值超过了临界的 F值，则附加的约束条件是无效的。特别注意：约束条件在理论上或许是有效的，但在具体的研究中却不一定。

1 这种方法可以很容易地推广到有多个线性约束条件的情形。

(续)

回到我们的例子中来，方程(14-8)相应的F值为：

$$F = \frac{(0.982\ 4 - 0.942\ 8)/1}{(1 - 0.982\ 4)/(15 - 3)} = 27.13 \tag{14-9}$$

在5%和1%显著水平下的临界的F值分别为4.75和9.33 (自由度分别为1和12)。显然，计算得到F值远远超过了F临界值，因此我们拒绝零假设：在研究期间内，就台湾经济而言，制造业中规模收益不变。有证据表明它是规模收益递增的。

注：方程(14-7)的R^2的实际值为0.249 3。但是，R^{*2}=0.942 8可直接与方程(14-3)给出的R^2=0.9824相比。我们知道，要比较两个R^2值，解释变量的必须是相同的。

方程(14-8)给出了F检验的一般步骤，它可以用于研究多个线性约束条件的情形。我们所要做的全部工作只不过是估计有约束和无约束回归方程，并利用方程 (14-8)所给出的F检验。如果F统计量是显著的，则我们拒绝有约束最小二乘回归，而使用无约束最小二乘回归。但是，如果F检验是不显著的，则我们选择RLS而不是无约束的OLS。

例14.2 对鸡肉需求

回到方程(10-15)和(10-16)所描述的对鸡肉的需求函数。这两个需求函数的差别在于方程(10-16)不包括解释变量猪肉价格和牛肉价格，但猪肉和牛肉可能是鸡肉的竞争产品。因此，方程(10-15)是无约束回归，而方程(10-16)则是有约束回归，因为我们明确假设这些变量对鸡肉的需求没有影响。回归方程 (10-15)的样本决定系数R^2=0.982 3，而RLS方程(10-16)的样本决定系数R^{*2}=0.980 1。因此，根据方程(14-8)得到：

$$F = \frac{(0.982\ 3 - 0.980\ 1)/2}{(1 - 0.9823)/(25 - 5)} = 1.118\ 6$$

(注意：m=2，n=23，k=5)服从分子和分母的自由度分别为2和18的F分布。观察F分布表发现，即使在10%的显著水平下，这一F值也是不显著的。显然，看来猪肉和牛肉的价格对鸡肉的需求并没有显著的影响。因此，RLS回归(10-16)比ULS回归(10-15)好。

在继续讨论之前，对于上述一例有几点需要指出：在第10章对多重共线性的讨论中曾指出，如果不正确地从模型中略去某个变量，则模型中剩余变量的系数将是有偏的 (参见第13章)。现在的问题是：究竟哪一个是正确的模型呢，是方程 (10-15)还是方程(10-16)？由于F检验表明可以接受RLS回归(10-16)，所以我们可以认为正确的模型是 (10-16)。因此，这个例子并不存在偏差问题。而且，正如第13章中所指出的那样，如果式(10-16)是正确的模型，而我们却把“多余的”变量(猪肉和牛肉的价格)包括进模型，即估计了模型 (10-15)，那么，OLS估计量仍然是无偏的，只不过它们的方差会变大。

14.2 动态经济模型：自回归模型和分布滞后模型

迄今为止，我们所讨论的模型都假设被解释变量 Y与解释变量X是同时期(contemporaneous)的，也即在同一时点。这一假设对截面数据是成立的，但对时间序列数据却不成立。因此，在消费支出对个人可支配收入 (PDI)的回归中(涉及到时间序列数据)，消费支出有可能取决于前一期和当期的PDI。也就是说，在Y和X之间可能存在滞后关系 (lagged)。

令，Y_t=t期的消费支出，X_t=t期的PDI，X_{t-1}= (t-1)期的PDI，X_{t-2}=(t-2)期的PDI。现在考虑

模型

$$Y_t=A+B_0X_t-B_1X_{t-1}+B_2X_{t-2}+u \tag{14-10}$$

模型(14-10)表明，由于滞后项X_{t-1}和X_{t-2}，消费支出与PDI不是同期的。类似式(14-10)这样的模型称为动态模型(dynamic models)(也就是说，涉及到跨时期变化)，因为解释变量每单位变化的影响分散在几个时期内，在方程(14-10)中是三个时期。

更专业地，我们把形如方程(14-10)这样的动态模型称为分布滞后模型(distributed lag models)，因为解释变量每单位变化的影响分布到了多个时期。为了进一步说明这一点，我们考虑以下假想的消费函数[1]：

$$Y_t=\text{常数}+0.4X_t+0.3X_{t-1}+0.2X_{t-2} \tag{14-11}$$

假定某人一年增加了薪金1 000美元，而且假定增加是“持久的”，就是薪金的增加长期不变。如果其消费函数形如方程(14-11)，则在第一年，他或她会增加消费支出400美元(0.4×1 000)，在第二年再增加支出300美元(0.3×1 000)，在第三年又支出200美元(0.2×1 000)。这样到第三年末，其消费支出将会增加(200+300+400)，即900美元；剩余的100美元用作储蓄。

与下面的消费函数相比：

$$Y_t=\text{常数}+0.9X_{t-1} \tag{14-12}$$

尽管，增加的1 000美元收入对消费的最终影响在两种情形下是相同的，但是在方程(14-12)中这种影响只滞后了一年，而在方程(14-11)中这种影响却分布到三年。因而，称形如式(14-11)的方程为分布滞后模型。这一点从图14-1中可以清楚地看到。

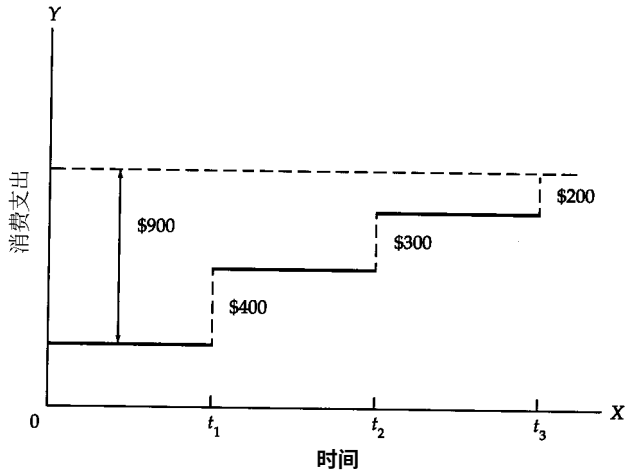


图14-1 分布滞后模型一例

很自然会提出这样一个问题：为什么会出现滞后现象呢？也就是说，为什么被解释变量会对滞后解释变量变化有反应呢？原因如下：

心理上的原因。由于习惯势力(惰性)，在价格下降或收入增加后，人们并不会立即改变他们的消费习惯。例如，那些因中彩而顷刻间成为百万富翁的人不可能立即改变他长期习惯了的生活方式，因为他还不知道该如何应付这突如其来的巨额财富，更不用说去应付那些闻风而来的金融计划者、新认的亲戚、税务律师等等。

技术上的原因。每次新一代的个人电脑(PC)一问世，市场上现用的PC机的价格都会急剧下降。因而，未来的计算机消费者就心存观望，花一定时间从容考察各种竞争品牌的特点与价格以后，才去购买。而且，他们不急于购买，还在希望价格下降得更低一点，或者希望出现更

1 这个例子取自Henri Theil，*Introduction to Econometrics*，Pretice-Hall，Englewood Cliffs,N.J.,1978,p.332.

有用的新产品。对汽车也如此。比方说，1999款汽车面市，则1998款汽车的价格会有较大幅度的下降。那些想要替换旧车的消费者可能会等待1999新款宣布上市后用较低的价格来购买98款的汽车。

制度上的原因。由于大多数集体议价协议都是多年期的契约，因此，工会工人们不得不等待现存契约到期后再商谈一个新的工资率，即使签约以来通货膨胀率一直以较大的幅度增加。类似地，一个职业球员也不得不等待他的契约期满后方可商谈另一个新契约，即使自签约以来，他的“能力”有了很大的提高。当然，也有球员尝试着重新商谈契约的，并且有些的确也获得了成功。

由于上述这些以及其他的一些原因，“滞后”在经济学中很重要。这一点清楚地反映在经济学的短期／长期分析方法上。在短期内，价格或者收入弹性的绝对值通常会小于长期内的价格或收入弹性值，因为解释变量变化之后，需要花时间去作必要的调整。

将方程(14-10)一般化，得到k期滞后分布模型：

$$Y_t=A+B_0X_t+B_1X_{t-1}+B_2X_{t-2}+\cdots+B_kX_{t-k}+u_t \tag{14-13}$$

在这个方程中，解释变量每单位变化的影响分布到了k个时期[1]。在方程(14-13)中，Y不仅对变量X本期值、前一期值的变化有反映，而且还对变量的前若干时期值的变化有反应。

在方程(14-13)中，系数B_0称为短期乘数(short-run multiplier)或者即期乘数(impact multiplier)，因为它给出了变量X每单位的变化所引起的同时期Y的平均变动量。如果X维持同样的变动水平，则(B_0+B_1)给出了下一时期Y变化的平均值，同理，$(B_0+B_1+B_2)$则给出了再下一时期Y变化的平均值，以此类推。这些系数之和称为中期乘数(interim, or intermediate multiplier). 最后，在k期后，得到

$$\sum_{i=0}^{k} B_i = B_0 + B_1 + B_2 + \cdots + B_k \tag{14-14}$$

称为长期乘数(long-run multiplier)或者总乘数(total multiplier)。因而，方程(14-11)消费函数的短期乘数为0.4，中期乘数为(0.4+0.3)=0.7，而长期乘数为(0.4+0.3+0.2)=0.9。长期地看(这里为三个时期)，平均而言，PDI每变化一单位将导致消费支出变化0.9单位。简言之，长期边际消费倾向(MPC)为0.9，而短期MPC仅为0.4，中期MPC为0.7。由于解释变量较前期的影响可能小于中期或近期的影响，所以一般而言，预期B_0大于B_1，而B_1大于B_2，如此等等。换言之，我们预期各个B值从本期往前递减，这对估计滞后分布模型会有所帮助。

分布滞后模型的估计

我们如何估计形如方程(14-13)的分布滞后模型呢？能否仍使用常用的OLS法？原则上说，是可以的，因为在重复抽样过程中，我们假设X_t是非随机的，或是固定的，因而X_{t-1}及其他所有X的滞后值也都是非随机的或固定的。所以，模型(14-13)本身并没有违背古典线性回归模型(CLRM)的任何假定。但是，在具体应用中还存在着一些实际问题。

(1) 对滞后的最大长度没有事先的指导或设想。

(2) 如果引入了太多期滞后变量，则自由度可能会是一个严重的问题。如果我们有20个观察值，引入10个滞后变量，则自由度仅为8——10个滞后变量就损失10个自由度——当期变量损失一个自由度，截距项也损失一个自由度。显然，随着自由度的减少，统计推断就会变得越来越不可靠。如果模型中的解释变量不止一个，并且每个解释变量都有其自己的分布滞后结构，则问题会变得越来越复杂。在这种情形下，我们会很快地“消耗”自由度——注意对每一个估

1 时期这一术语是广义的，它可以是一天、一个星期、一个月、一年或者是任何合适的时间段。

计系数，都将失去一个自由度。

(3) 在大样本情形下，虽然无需过多考虑自由度，但可能会遇到多重共线性问题，因为大多数经济变量的连续值是趋于相关的，有时相关程度还很高。在第 10章我们曾指出，多重共线性会导致估计不准确，也就是说，估计系数的标准差会变大。结果，根据常规计算的 t值，往往认为滞后系数是统计不显著的。产生的另一个问题是，有的时候，滞后项系数的符号会出现正负交错的情况，这使得我们很难解释这些系数，从下面的例子将会看到这一点。

例14.3 具体一例：圣·路易斯 (St.Louis) 模型

为了确定名义国民生产总值 (GNP) 是用货币供给(货币主义)，还是用政府支出(凯恩斯主义)来解释，圣·路易斯联邦储备银行建立了一个模型，这就是著名的圣·路易斯模型。这个模型的其中一个版本是

$$\dot{Y}_t = \text{常数} + \sum_{i=0}^{4} A_i \dot{M}_{t-i} + \sum_{i=0}^{4} B_i \dot{E}_{t-i} + u_t \qquad (14\text{-}15)$$

其中，

$\dot{Y}_t$ ——t时期的名义GNP增长率

$\dot{M}_t$ ——t时期的货币供给(M_1)增长率

$\dot{E}_t$ ——t时期充分就业或高就业下政府支出的增长率

样本数据是从1953年第一季度到1976年第四季度的季度数据，$\dot{M}$和$\dot{E}$各使用了四个滞后值。回归结果如下：[1] 为了阅读的方便，用表格的形式给出回归结果(表14-1)

表14-1 圣·路易斯模型

系数	估计		系数	估计	
A_0	0.40	(2.96)*	B_0	0.08	(2.26)*
A_1	0.41	(5.26)*	B_1	0.06	(2.52)*
A_2	0.25	(2.14)*	B_2	0.00	(0.02)
A_3	0.06	(0.71)	B_3	−0.06	(−2.20)
A_4	−0.05	(−0.37)	B_4	−0.07	(−1.83)*
$\sum_{i=0}^{4} A_i$	1.06	(5.59)*	$\sum_{i=0}^{4} B_i$	0.01	(0.40)
	R^2=0.40		d=1.78		

注：括号内的数字为t值。

* 显著水平为5%(单边)。原文中并没有给出截距项。

对表14-1所给的结果有几点需要注意：

(1) 根据t检验，并非所有的滞后系数均是统计显著的。但是我们无法得知确实是不显著，还是由于多重共线性的影响。

(2) $\dot{M}$的第四个滞后变量的符号为负，这在经济学上很难解释，因为货币供给的其他滞后系数都对$\dot{Y}$产生正向影响。但这一负值是统计不显著的，尽管我们不知道这是否是多重共线性

1 这些结果，除了符号略有改动外，均取自Keith M. Carlson的“Does the St.Louis Equation Now Believe in Fiscal Policy”，Review，Federal Bank of St.Louis，vol.60，no.2，February,1978，Table IV，p.17。注：

$\sum_{i=0}^{4} A_i \dot{M}_{t-i} = A_0 \dot{M}_t + A_1 \dot{M}_{t-1} + A_2 \dot{M}_{t-2} + A_3 \dot{M}_{t-3} + A_4 \dot{M}_{t-4}$

$\sum_{i=0}^{4} B_i \dot{E}_{t-i}$ 可类似得到。

所致。$\dot{E}$的第三和第四个滞后变量不仅为负而且也是统计显著的。同样，经济学难以解释这一负值，因为无法解释为什么政府支出增长率在第三和第四期有负向影响而在前两个时期的影响却是正向的？

(3) $\dot{M}$每单位变化，对名义GNP的短期影响为0.40，而长期影响为1.06(即各个A系数加总)，并且它是统计显著的。对此的解释是，货币供给持续增长 1%，名义GNP在五个季度的增长率约为1%。类似地，政府支出增长率每增长 1%，对名义GNP的短期影响为0.08，并且是统计显著的，但其长期影响仅为0.01(B系数的加总)而且还是统计不显著的。

这隐含地表明，货币供给增长率的变化对 GNP增长率的变化有着持久的影响 (几乎是一对一)，但政府支出增长率的变化却不是如此。简言之，圣·路易斯模型支持了货币主义的理论，所以圣·路易斯模型常常被称作是货币主义模型。

从统计学的观点来看，一个明显的问题是为什么圣·路易斯模型的每个解释变量仅包括四期滞后值呢？一些不显著的系数是否是由于多重共线性所致？在没有检验初始数据和确定引入更多个滞后项会发生什么结果之前，我们无法回答这些问题。但是，读者可以想象，沿着这个方向探索下去不会有丰硕的成果，因为如果引入了多个滞后项，难免会产生多重共线性的问题。显然，我们需要一种方法，它不仅能避免多重共线性问题，而且还能告诉我们模型中可以包括多少个滞后变量。

14.3　用于分布滞后模型的夸克方法[1]

一种既能减少分布滞后模型中滞后项个数又能解决多重共线性问题的创造性方法就是夸克所采用的适应性预期(adaptive expectations)模型，和部分或存量调整模型 (the partial or stock adjustment models)。这里，我们不去深究这些模型的技术细节，[2] 所有这些模型的一个显著特点是，形如方程(14-13)的分布滞后模型可以简化成如下“简单”形式：[3]

$$Y_t=C_1+C_2X_t+C_3Y_{t-1}+v_t \tag{14-16}$$

其中v为误差项。这种模型称为自回归模型(autoregressive model)，因为被解释变量的滞后值作为解释变量出现在方程的右边。在方程 (14-13)中，我们需要对截距项，解释变量的当期及k期滞后项进行估计。因此，如果 k=15，我们将共估计 17个参数，自由度的损失是相当大的，尤其是对容量不太大的样本。但是在回归方程 (14-16)中，只需估计三个未知变量，截距项及两个斜率系数，大大地减少了自由度的损失。也就是说，回归方程 (14-13)中的所有滞后项都由单独的一个Y的滞后值来代替。

当然，没有“免费的午餐”。在将模型(14-13)中估计的参数个数减少到只有三个的过程中，也产生了一些新的问题。首先，由于Y_t是随机的，Y_{t-1}也是随机的。因此，要用OLS来估计模型，必须确保误差项v_t与滞后变量Y_{t-1}不相关；否则，可以证明，OLS估计量不仅是有偏的，而且也是不一致的。然而，如果 v_t和Y_{t-1}不相关，则可以证明，OLS估计量是有偏的 (对小样本而言)，

1　参见L. M. Koyck, *Distributed Lags and Investment Analysis*，North-Holland，Amsterdam，1954；P. Cagan,“The Monetary Dynamics of Hyper Inflation,” in M.Friedman(ed.,)，University of Chicago，Chicago，1956(对适应性预期模型)；Marc Nerlove，“Distributed Lags and Demand for Agricultural and Other Commodities , Handbook No.141, U.S. Department of Agricultural, June 1958(对部分或存量调整模型)。

2　有关技术细节可参见：Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，Chap.16。

3　事实上，方程(14-13)中的滞后项的个数可以远远地超过 k个，从技术上是无限的。顺便指出，注意夸克模型和其他两个模型都隐含地假定了 B系数的值是递减的；也就是说，从数值上讲，B_0会大于B_1，而后者又会大于B_2，如此等等。这是一个可行的假设，因为较远的过去不可能和当前具有相同的影响。

但是这一偏差会随着样本容量的增大而逐渐消失。也就是说，对大样本而言， OLS估计量是一致估计量(从理论上说是渐进的一致估计量)。其次，如果v_t是序列相关的(例如，它服从一阶马尔可夫过程，$v_t=\rho v_{t-1}+w_t$，其中$-1\leq\rho\leq1$，并且误差项ω_t满足OLS假定)，则OLS估计量是有偏的和不一致的，传统的t检验和F检验也是无效的。因此在形如方程 (14-16)的自回归模型中，非常重要的一点是我们确定误差项 v_t是否服从一阶马尔可夫或者 AR (1)过程。第三，在第12章中曾指出，在自回归模型中，传统的杜宾 -瓦尔森d检验不再适用了。在这种情况下，可以利用习题12.1中的杜宾h统计量检验一阶自回归，或者我们可以进行趋势检验。

在继续说明模型 (14-16)之前，需要指出的是：变量 X_t的系数C_2给出了单位X_t的变化对Y_t的短期影响，而$C_2/(1-C_3)$则给出了单位X_t的(持续)变化对于Y_t的长期影响；这等同于将模型中所有B系数加总，如方程 (14-14)所示[1]。换句话说，回归 (14-16)中Y的滞后项起着模型 (14-13)中所有滞后X项的作用。

例14.4 调整后的基础货币增长率对名义GNP增长率的影响，美国，1960~1988年

为了弄清楚名义GNP增长率与调整后的基础货币 (AMB) [2]增长率之间的关系，约瑟夫H.哈斯拉格(Joseph H. Haslag)和苏格特E.海因(Scott E. Hein) [3]得到了以下回归结果：

$$\hat{Y}_t=0.004+0.238\text{AMB}_{t-1}+0.759Y_{t-1} \qquad \text{杜宾的}h=3.35$$
$$\text{se}=(0.004)\quad(0.067)\quad(0.054) \qquad (14\text{-}17)$$
$$t=(1.000)\quad(3.552)\quad(14.056)$$

(注意：作者并没有给出R^2值。)

注：变量上面一点表示了增长率。

注意：这里，Haslag和Hein使用了滞后一期的AMB(这里是一年)作为解释变量，而没有使用当期的AMB，但这么做并不会产生任何问题，因为 AMB在很大程度上是由联邦储备系统决定的。除此之外，如果AMB_t是非随机的，则AMB_{t-1}也是非随机的，这满足标准的 CLRM假定。现在，我们来解释模型 (14-17)。

从方程(14-17)可知， AMB的短期影响是0.238；也即AMB每变化一个百分点，平均而言，将导致名义GNP变动0.238个百分点。这一影响看来是统计显著的，因为计算的 t值是显著的。但是，长期影响为

$$\frac{0.238}{(1-0.759)}=0.988 \qquad (14\text{-}18)$$

接近于1。因此，从长期来看， AMB(持续的)每变动1%，将导致名义 GNP变动大约一个百分点；也就是说， AMB增长率与名义GNP增长率之间存在着一对一的关系。

模型(14-17)的惟一问题是估计的h值是统计显著的。正如在练习题 12.16中所指出的那样，

1 详细的讨论参见 ，Damodar N. Gujarati，Basic Econometrics，3d ed., McGraw-Hill，New York，1995，Chap.16。

2 基础货币(MB)，有时也称为高能货币，在美国它包括现金和商业银行总储备金。 AMB考虑了联邦储备银行所要求的储备金率的变化；在美国，所有的商业银行都要求保留与顾客储蓄相对应的一定数量的现金或者现金等价物。储备金率是现金和现金等价物与总储蓄的比例。联邦储备系统不时地改变这一比例以达到一定的政策目标，例如抑制通货膨胀或者控制利息率等等。

3 参见Joseph H. Haslag和Scott E. Hein，" Reserve Requirements,the Monetary Base and Economic Activity "，Economic Review，Federal Reserve Bank of Dallas，March 1989，p.13。回归结果是以适应性模型(14-15)的形式给出。

对大样本而言，h统计量服从标准正态分布。因此，在5%显著水平下，双边检验的Z临界值(标准正态)为1.96，而在1%显著水平下，双边检验Z临界值为2.58。由于观察到h值为3.35，超过了这些临界值，看来回归(14-17)中的残差是自相关的，因此模型(14-17)所给出的结果需谨慎对待。但是注意：h统计量是一个大样本统计量，而模型(14-17)的样本容量为29，并不算很大。但无论如何，方程(14-17)只是用于教学目的，用以说明用夸克模型，适应性预期模型和存量调整模型估计分布滞后模型的内在机制。

例14.5 保证金与股票市场不稳定性

为了估算保证金(这限制了经纪人和交易商提供给顾客的信贷数量)的短期和长期影响，吉卡斯 A.哈都维利斯(Gikas A. Hardouvelis)[1]根据从1931年12月到1987年12月，共673个月，股票的月度数据(包括标准普尔(S&P)指数)，得到以下回归结果：

$$\hat{\sigma}_t = 0.112 - 0.112 m_t + 0.186\sigma_{t-1}$$

$$\text{se} = (0.015)\quad (0.024)\quad (\ )^*\quad R^2 = 0.44 \qquad (14\text{-}19)$$

其中，σ_t是从$(t-11)$期到t期计算的月度超额名义股票收益率(名义收益率减去前一个月月末一个月期的债券收益率)的标准方差，这是不稳定性的一个度量指标；m_t是从$(t-11)$期到t期的官方平均保证金；括号中的数字是异方差和自相关校正后的标准差。不幸的是，Hardouvelis并没有给出σ_{t-1}系数的标准差，也没有给出h统计量。但作者已经在他的结果中就自相关进行了校正。

* 作者并没有给出标准差。

与预期相同，保证金系数的符号为负，表明当保证金增加时，在股票市场上的投机活动会减少，从而减少了不稳定性。－0.112表明如果保证金增加一个百分点，则S&P股票的不稳定性将会减少0.11个百分点。当然，这是短期影响。长期影响是

$$-\frac{0.112}{(1-0.186)} \approx -0.138$$

显然高于(绝对值)短期影响，但是高出的幅度并不大。

动态建模的方法很多，并且处理这类模型各种新的经济计量技术也不断出现。上述的讨论仅仅是向读者粗略地介绍了动态建模的概念。有关动态建模更详细的讨论可参阅有关的参考书。[2]

14.4 解释变量是虚拟变量的情形

在第9章中，我们详细地讨论一个或多个解释变量是虚拟变量(或定性变量)，但应变量是定量变量的回归模型。在本节考虑的模型中，被解释变量也是虚拟变量；而解释变量或是虚拟变量，或是定量变量，或是既有虚拟变量又有定量变量。

假设我们要研究成年男性劳动力参与状况，这里解释变量包括失业率、平均工资率、家庭收入、教育水平等。则某个人要么是劳动力，要么不是劳动力。因此，被解释变量，劳动力参与率，可能取两个值：1，如果该人是劳动力；或0，如果该人不是劳动力。我们来看另一个例子，假设我们想确定一个国家是否属于国际货币基金组织(IMF)，它可以是好几个因素的函数。那么，某个国家要么属于IMF，要么不属于IMF。因此，被解释变量，IMF成员，就是一个虚拟

1 参见Gikas A. Hardouvelis，“Margin Requirements and Stock Market Volatility”，*Quarterly Review*，vol.,3, no.2, Summer 1988，Table 4，p.86, and footnote21, p.88。

2 A. C. Harvey，*The Econometric Analysis of Time Series*, 2d ed., MIT，Cambrige，Mass，1990。需提醒初学者的是，这本书难度较大。

变量。如果该国属于IMF，则取值为1；否则取值为0。再来看另一个例子，美国假释局正在考虑假释一个未到期的罪犯。假释变量是一个虚拟变量，因为该局可能同意也可能不同意假释；同意假释的决定则是根据罪犯犯罪的类型，罪犯服刑期间的表现，是否患有精神病等若干因素。

读者可以举出许多这样的例子。这些例子的一个显著特征是：应变量只有“是”或“不是”两种反应，也就是说，它在性质上是“二分变量”(dichotomous)。如何估计这些模型呢？能否直接使用普通最小二乘法呢？我们通过一个具体例子来回答这类及其相关问题。

例14.6 个性化教育系统(PSI)对于中级宏观经济学课程成绩的影响

表14-2给出了以下一些数据资料：被解释变量，中级宏观经济学考试成绩 (Y)，并且，若成绩为A，则Y=1；若成绩为B或C，则Y=0，解释变量，GPA=入学的平均成绩，TUCE=新生入学后宏观经济学考试成绩，如果使用了个性化教育系统(personalized system of instructions)这一新方法，则PSI=1，否则为0。

表14-2 32个学生的平均成绩及相关数据

观察号	Y	GPA	TUCE	PSI	观察号	Y	GPA	TUCE	PSI
1	0	2.66	20	0	17	0	2.75	25	0
2	0	2.89	22	0	18	0	2.83	19	0
3	0	3.28	24	0	19	0	3.12	23	1
4	0	2.92	12	0	20	1	3.16	25	1
5	1	4.00	21	0	21	0	2.06	22	1
6	0	2.86	17	0	22	1	3.62	28	1
7	0	2.76	17	0	23	0	2.89	14	1
8	0	2.87	21	0	24	0	3.51	26	1
9	0	3.03	25	0	25	1	3.54	24	1
10	1	3.92	29	0	26	1	2.83	27	1
11	0	2.63	20	0	27	1	3.39	17	1
12	0	3.32	23	0	28	0	2.67	24	1
13	0	3.57	23	0	29	1	3.65	21	1
14	1	3.26	25	0	30	1	4.00	23	1
15	0	3.53	26	0	31	0	3.10	21	1
16	0	2.74	19	0	32	1	2.39	19	1

资料来源：L. Spector和M. Mazzeo，“Probit Analysis and Economic Education，”*Journal of Economic Education*，vol.11，1980，p.37-44。

Spector和Mazzeo研究的主要目标是评估新教学方法(PSI)对最终成绩的效果。

假定对如下模型使用OLS法拟合表14-2所给数据：

$$Y_i=B_1+B_2\text{GPA}_i+B_3\text{TUCE}+B_4\text{PSI}+u_i \tag{14-20}$$

其中应变量的取值为1或0，取决于是否获得A级成绩。这样得到的最终模型，也即方程(14-20)称为线性概率模型(LPM)，因为E(Y_i)可以解释为条件概率(参见第2章)，也即在给定解释变量取值的条件下，事件 Y(即某个学生得A)发生的概率[1]。因此，我们方程(14-20)可以表示为两个等价的形式：

$$E\ (Y=1|\text{GPA, TUCE, PSI})=B_1+B_2\text{GPA}_i+B_3\text{TUCE}_i+B_4\text{PSI}_i \tag{14-21}$$[2]

1 有关证明参见Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，Chap.15。

2 注意：$E\ (Y|\text{GPA, TUCE, PSI})=1\cdot P\ (Y=1|\text{GPA, TUCE, PSI})+0.P(Y=0|\text{GPA, TUCE, PSI})=P\ (Y=1|\text{GPA, TUCE, PSI})$

(续)

或

$$P\ (Y=1|\text{GPA, TUCE, PSI})=P_i=B_1+B_2\text{GPA}_i+B_3\text{TUCE}_i+B_4\text{PSI}_i \tag{14-22}$$

其中P代表概率。则

$$\hat{P}_i = b_1 + b_2\text{GPA}_i + b_3\text{TUCE}_i + b_4\text{PSI}_i \tag{14-23}$$

给出了某学生在给定解释变量取值的条件下得A的概率的估计值。

一个重要的问题是：迄今为止，我们只考虑解释变量为数量型变量的线性回归模型。而模型(14-20)并非如此，因为Y是一个虚拟变量，则使用OLS法估计方程(14-20)会不会有什么问题呢？的确会有问题。可以证明，使用OLS估计模型(14-20)带来以下问题：[1]

(1) 模型(14-20)中的误差项u_i不服从正态分布，事实上它服从二项(概率)分布。如果仅仅只是估计参数的话，并不需要正态分布的假设，但对假设检验却是需要的。但是，在实际中，如果样本容量足够大，模型(14-20)中误差项服从二项分布并不会造成很大的障碍。这是因为，统计理论表明，随着样本容量的增加，二项分布趋于正态分布。

(2) OLS估计的另一个问题是误差项u_i是异方差的。但在实践中这也不是一个很严重的问题，因为我们可以通过适当的变换使误差项成为同方差的。

(3) 模型(14-20)的真正问题是由于它给出的是事件Y发生的概率(例如获得A这一事件)，所以取值必须介于0和1之间(为什么？)。但是根据回归方程(14-23)获得这一概率时，并不能保证所估计的P_i一定会介于0、1之间。例如，如果估计的某个P_i为负，则它就没有实际意义，同样如果大于1，也没有实际意义。

(4) 模型(14-20)或者其等价方程(14-21)的另一问题是它假定每单位解释变量变化的概率变化率是一个常数，由斜率值给出。因而，B_2表示了GPA增加一单位，则获得A的概率增加了B_2个单位，而无论初始GPA的值是多少，这在实践中可能是一个不切实际的假设。事实上，如果考虑到“收益递减”规律，我们可能会预期获得A的概率将会以一个递减的速率增加。

正是由于这些原因，尤其是第三个原因，所以通常不用OLS法估计形如方程(14-20)的模型。但是，在讨论其他估计方法之前，我们用表14-2提供的数据来说明模型(14-20)，看一看在这个例子中是否存在问题(3)。对模型(14-20)直接用OLS法进行估计，得到如下结果：

$$\begin{aligned}\hat{P}_i(Y=1) &= -1.4980 + 0.4639\text{GPA}_i + 0.011\text{TUCE}_i + 0.3786\text{PSI}_i \\ \text{se} &= (0.523\ 9)\quad (0.161\ 96)\quad (0.019)\quad (0.139\ 17) \\ t &= (-2.859\ 4)\quad (2.846\ 1)\quad (0.538\ 7)\quad (2.720\ 0)\quad R^2=0.415\ 9\end{aligned} \tag{14-24}$$

(注意：这些结果并没有校正异方差。)

对模型解释如下：在其它情况不变的条件下，GPA每增加1分，平均而言，获得A级的概率将增加0.46，而无论初始的GPA为多少。类似地，对于那些已经接受了PSI的学生来说，他们获得A级的概率将增加0.38。显然，在这个例子中TUCE对获得A级的概率并没有明显的影响。因此，这一新的教学方法看来是有效的。

现在如果将方程(14-24)右边解释变量的值带入，则会得到获得A级概率的实际估计值。例如，如果是第22个学生，将对应的值带入方程，得到：

$$-1.540\ 2+0.477\ 8\ (3.62)+0.010\ 7\ (28)+0.3731\ (1)=0.853\ 54$$

也即该学生获得A级的概率为0.86。如果取第32个学生，进行类似的计算，求得概率为0.1781。

1 有关证明参见Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，Chap.15 p.542-544。

但是，如果取第一个学生，并进行同样的计算，你会发现概率为- 0.055 3，是个负的概率！实际上若表中的每个学生都进行同样的计算，将会发现有五个学生得到 A级的概率为负！也就是说，有15%的情况，概率值是负值。正是由于这个原因，所以我们并不提倡用 LPM方法来估计应变量具有“二分性的”的模型。

14.5 分对数模型

还有其他的抉择模型(alternatives)吗？在文献中讨论得比较多的有另外两种方法，分对数模型和概率模型。由于两个模型给出的结果大致相同，因此我们仅仅讨论分对数模型，因为相对来说它在数学上比较简单[1]。我们仍用上面的一个例子来说明分对数模型。这里仍假定 P_i为获得A等级成绩的概率。现考虑如下模型：

$$\ln\left(\frac{P_i}{1-P_i}\right)=B_1+B_2\text{GPA}_i+B_3\text{TUCE}_i+B_4\text{PSI}_i+u_i \tag{14-25}$$

由于P_i是获得A等级成绩的概率，则$(1-P_i)$则是没有获得A级的概率。比例$P_i/(1-P_i)$称为差别比率(odds ration)。该差别比率的自然对数称作分对数，因此，模型 (14-25)称为分对数模型。[2]根据分对数模型可知差别比率的对数是解释变量的线性函数。在我们这个例子中解释变量为GPA，TUCE和PSI。在模型(14-25)中，斜率系数，比如说，B_2给出了GPA每单位的变化，所引起的差别比率对数的改变量。注意，与 LPM不同，分对数模型并不直接给出概率值。稍后我们将说明如何计算概率。

分对数模型有什么特点呢？首先，模型的数学形式保证了从分对数模型中所估计出概率的值总是介于0、1之间。其次，与LPM不同，在这个模型中，解释变量每变化一个单位，获得 A级的概率不是线性增加的 (即不是以一个固定值增加)，事实上，当解释变量变得越来越小时，概率则以越来越小的速率接近于零，当解释变量变得越来越大时，概率则以越来越小的速率接近于1，如图14-2所示。

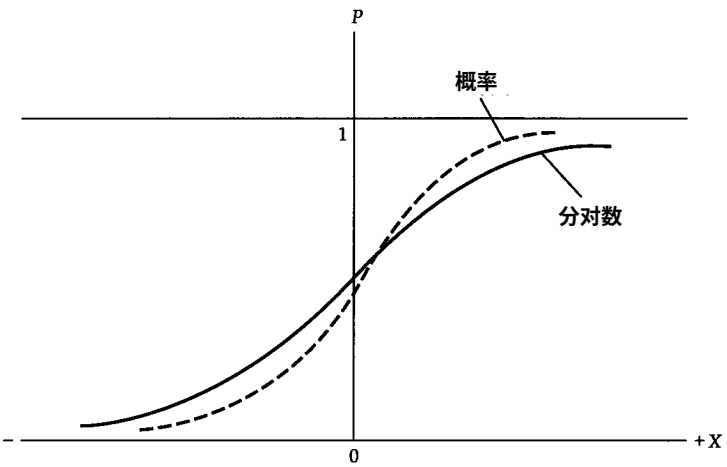


图14-2 分对数和概率累积分布函数

1 有关概率模型和分对数模型的讨论，参见 Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，Chap.16。

2 分对数模型名称来源于逻辑斯蒂概率分布函数，有关逻辑斯蒂概率分布函数的讨论参见 Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，p.554-556。

14.5.1 分对数模型的估计：个体数据

分对数模型(14-25)与第8章讨论过的半对数模型类似。看起来可用常用的 OLS方法进行估计。不幸的是，情况并非如此，因为要用 OLS法估计方程(14-25)，必须知道被解释变量 $\ln P_i/(1-P_i)$的值，而显然它是未知的。事实上，我们的最终目标是确定在给定解释变量的值的情形下P是多少。那么如何估计形如方程 (14-25)这样的模型呢？答案取决于用于分析数据的类型。如果我们有个体观察数据，则可以使用最大似然法 (ML)来估计模型。如果是分组数据 (如表14-3)，则可以使用OLS法。由于GPA的数据是以每个学生的形式给出，所以可以使用 ML方法来估计模型(14-25)中的参数。这里，我们不打算详细讨论这一方法，它超出本书的范围，但是需要指出的是有许多计算机软件包可用 ML法估计分对数模型。对于本例，我们用的是 SHAZAM软件。我们首先给出ML估计的结果。稍后我们给出当数据类型是分组数据时，如何使用 OLS法。

根据SHAZAM程序，得到模型(14-25)的回归结果如下：[1]

$$\ln\frac{P_i}{1-P_i} = -13.021 + 2.826\,1\mathrm{GPA}_i + 0.095\mathrm{TUCE}_i + 2.378\,7\mathrm{PSI}_i$$
$$\mathrm{se} = (4.931\,0)\quad (1.262\,9)\quad (0.142\,155)\quad (1.064\,5)$$
$$t = (-2.640\,7)\quad (2.237\,8)\quad (0.722\,6)\quad (2.234\,5) \qquad (14\text{-}26)$$
$$\text{Maddala}R^2 = 0.382\,1$$

对回归结果解释如下：在其他条件不变的情况下， GPA上升一单位，平均而言，获得 A级的分对数(或差别比率的对数)将上升2.8个单位。[2] 类似地，在其它条件不变的情况下，平均而言，对于那些接受 PSI的学生来说，获得 A级的差别比率(对数)上升2.38单位。显然，TUCE对GPA有可检验到的影响。

我们如何计算获得A级的实际概率呢？为了求这个概率值，我们首先说明如何计算任意一个学生的分对数值。将表 14-2中对应每个学生的解释变量的值代入方程 (14-26)，获得$\ln P_i/(1-P_i)$的值。例如，对于第一个学生而言，计算的分对数值为

$$\ln\frac{P_i}{1-P_i} = -3.600\,7 \qquad (14\text{-}27)$$

现在若取方程(14-27)的反对数，得：

$$\frac{P_i}{1-P_i} = e\{-3.617\,4\} \qquad (14\text{-}28)$$

其中$e\{\}$表示e的$\{\}$次幂，这里，$\{\}$内的值为−3.6174。对于方程(14-28)做简单的代数变换得到：

$$P_i = \frac{e\{-3.6174\}}{1+e\{-3.617\,4\}} = 0.026\,6 \qquad (14\text{-}29)$$

注：$e\{-3.617\,4\} \approx 0.026\,8$ (14-30)

也就是说，对于给定的解释变量值，该学生获得 A级的概率为3%。若取第30个学生，并进行类似的计算，他或者她的分对数值为 2.850 4，根据这个值可以估计得到概率为 0.945 34。这意味着该生获得A级的概率为95%。作为一般规则分对数值越高，则获得 A级的可能性就越大，因而获得A级的概率也就越大。在习题 14.17中，要求其他30个学生获得A级的概率，这些学生

---

1 一些技术要点：在以下回归中所给出的标准差和t值是渐近的，也就是说，它们是对大样本而言的。其次，对于分对数模型而言，通常的 R^2不再有意义。有其它几个衡量指标，例如 Maddala，Cragg-Uhler和McFaddenR^2。在本例中，MaddalaR^2和McFaddenR^2两者的值大致相同的，但Cragg-UhlerR^2值为0.536 4。

2 特别需要注意的是，斜率系数给出的是对于一单位解释变量的变化，差别比率对数的变化而不是概率本身的变化。后一变化计算起来有些困难。有关细节可以参阅 Damodar N. Gujarati，“*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，p.559-560。

的成绩在表14-2中给出。

14.5.2 分对数估计：分组数据

例14.7 住房所有权的分对数估计

假定我们要研究住房所有权对家庭收入的函数。现有表 14-3中给出的假想数据。共有580个家庭，但与表14-2不同的是，并没有有关单个家庭收入与住房所有权状况的数据。实际上，是把这580个家庭分成10个收入等级，有每个等级的家庭总数及实际拥有住房的家庭的数据，我们把这样的数据称为**分组数据**(grouped data)。而表14-2中的数据则是微观或个体数据。现在取 n_i 对 N_i 的比值，得：

$$p_i = \frac{n_i}{N_i} \tag{14-31}$$

也就是相对频率(参见第2章)，它可用作真实概率 P_i 的估计量，尤其当 N_i 非常大时。[1] 因此在分对数模型中：

$$\ln \frac{P_i}{1-P_i} = B_1 + B_2 X_i + u_i \tag{14-32}$$

P_i=拥有住房的概率，X_i=收入，如果把从方程(14-31)得到 P_i 带到方程(14-32)，则方程左边就很容易计算得到，这样 OLS就可以直接用于方程(14-32)。根据表14-3所给的数据，并利用表第4列中给出的 p_i，得到方程(14-32)的回归结果如下：

$$\begin{aligned} \ln \frac{\hat{p}_i}{1-\hat{p}_i} &= -3.243\,8 + 0.079\,2 X_i \\ \text{se} &= (0.170\,8) \quad (0.004\,1) \\ t &= (-18.992) \quad (19.317) \qquad R^2 = 0.979\,1 \end{aligned} \tag{14-33}$$

表14-3 X_i(收入)，N_i(收入为 X_i 的家庭数)和 n_i(拥有住房的家庭数)

X (千美元)	N_i	n_i	$p_i = \frac{n_i}{N_i}$
(1)	(2)	(3)	(4)
26	40	8	0.20
28	50	12	0.24
30	60	18	0.30
33	80	28	0.35
35	100	45	0.45
40	70	36	0.51
45	65	39	0.60
50	50	33	0.66
55	40	30	0.75
60	25	20	0.80
	580	269	

正如回归结果所表明的那样，如果收入增加了一单位 (这里是1 000美元)，拥有住房的差别比率的对数将增加 0.08个单位。当然，实际上我们可以计算任何给定收入水平下拥有住房的概率。例如，令 X=26，得：

1 回忆我们在第2章关于概率的讨论，事件的概率是随着样本无限增大的相对频率的极限。

(续)

$$\ln \frac{p_i}{1-p_i} = -1.184\ 6$$

因此，从方程(14-29)中可得：

$$p_i(\text{给定}X=26)=0.234\ 2 \qquad (14\text{-}34)$$

这里，实际概率为0.20。其它概率可类似计算(参见习题14.18)。

方程(14-33)有一个技术要点：这个回归存在异方差问题。对于以分组数据为基础的分对数模型而言的确如此。但是这个问题很容易解决，详细内容参阅有关参考书。[1]

掌握了分对数模型，读者可以很容易地处理被解释变量为虚拟变量或二元变量的模型。事实上，我们甚至可以对被解释变量和解释变量都是虚拟变量的模型进行回归。不仅如此，而且我们还可以考虑那些被解释变量为三元变量(例如属于共和党或者民主党或者第三党)，甚至被解释变量分更多类的模型。这类模型称为多值回归模型(multinominal regression)，但它已超出本书范围。[2]

下面我们给出分对数模型的一个有趣应用。

例14.8　预测银行破产

根据1982年12月到1984年12月间6869次例行检查所得出的银行例行报告数据，罗伯特·阿维莱(Robert Avery)和特伦斯·贝尔顿(Terrence Belton)[3]估计了一个风险指数(也就是一个分对数函数)用来预测银行倒闭：如果在例行检查后的一年之内破产了的话，则认为该银行已经倒闭了。这些结果在表14-4中给出。

表14-4　分对数模型：预测银行倒闭

解释变量	系数	t值
常数	-2.420	3.07
KTA	-0.501	-4.89
PD090MA	0.428	5.16
LNNACCA	4.310	4.31
RENEGA	0.629	1.07
NCOFSA	0.223	1.60
NETINCA	0.331	2.68

注：　KTA——原始资本对总资产的百分比。
PD090MA——超过90天的贷款对总资产的百分比。
LNNACCA——非增值贷款对总资产的百分比。
RENEGA——重新议定贷款对于总资本的百分比。
NCOFSA——(每年)净损耗贷款对总资产的百分比。
NETINCA——(每年)净收入对总资产的百分比。

数据来源：Robert Avery和Terrence Belton，“A Comparison of Risk-Based Capital and Risk-Based Deposit Insurance，”*Economic Review*，Federal Reserve Bank of Cleveland，1987，fourth quarter。

1　Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，p.558-561。

2　关于逻辑斯蒂模型的一般讨论，参见　David M. Hosmer，Stanley Lemesshow，*Applied Logistic Regression*，Wiley，New York，1988。

3　结果来自Robert Avery和Terrence Belton，“A Comparison of Risk-Based Capital and Risk-Based Deposit Insurance，”*Economic Review*，Federal Reserve Bank of Cleveland，1987，fourth quarter，p.20-22。

回归结果表明，在其他情况不变的条件下，如果KTA，原始资本(也就是股东股票)对总资产的比率，每增加一个百分点，则银行倒闭的差别比率的对数增加0.501，这是一个比较合理的结果。类似地，如果LNNACCA，非增值贷款对总资产的百分比，每增加一个百分点，则银行倒闭差别比率的对数增加4.310。其他系数也可以类似地加以解释。RENEGA和NCOFSA都是统计不显著的，并且NCOFSA的符号也不正确。(为什么？)

Robert Avery和Terrence Belton得出的结论是：

尽管模型的总体拟合表明，预测银行倒闭是很困难的，平均而言，预测样本中破产银行破产的概率为0.24，这一数字比预测样本中非破产银行破产的概率高69倍。因此，这一模型显然具有区分高风险银行和低风险银行的能力。

14.6 伪回归现象

有时，从时间序列模型得到的回归结果从表面上看是很好的，但进一步细究则发现结果是值得怀疑的。我们通过一个具体例子来说明伪回归(spurious regression)现象。表14-5给出了从1970年第一季度到1991年第四季度美国国内生产总值(GDP)，个人可支配收入(PDI)，个人消费支出(PCE)，利润、红利等数据，所有的数据的单位都是10亿美元(1987年美元价)，共有88个观察值。这里我们仅考虑PCE和PDI，表中给出的其它数据将在本章习题中用到。

利用表14-5提供的数据并作PCI对PDI的回归，回归结果如下：

$$\mathrm{PCF}_t = -171.441\ 2 + 0.967\ 2\mathrm{PDI}_t \qquad R^2 = 0.994\ 0;\ d = 0.531\ 6$$

$$t = (-7.480\ 9) \quad (119.871\ 1) \tag{14-35}$$

回归结果看似令人“难以置信”；R^2值非常高，PDI的t值也非常高，PDI的边际倾向(MPC)为正并且很高。惟一的缺陷是d值较低。Granger和Newbold指出，若$R^2 > d$，则很可能存在伪回归现象，也就是说，实际上，PCE和PDI之间可能并不存在任何有意义的关系。[1]

为什么方程(14-35)中的回归结果可能是虚假的呢？为了便于理解，我们引入平稳时间序列(stationary time series)的概念。我们首先根据表14-5中PCE和PDI的数据作散点图，如图14-3所示。从图14-3可以看出PCE和PDI这两个时间序列在样本期间都呈上升趋势。这样的图形通常表明该时间序列可能是不平稳的。这意味着什么呢？宽泛地说，随机过程称为平稳的，如果其均值和方差在长时期内为常数，并且两个不同时期的协方差仅仅依赖于这两个时期的“距离”或时滞而不依赖于计算协方差的实际时点。[2]

表14-5 1970年第一季度到1991年第四季度美国的宏观经济数据

季度	GDP	PDI	PCE	利润	红利
1970-1	2 872.8	1 990.6	1 800.5	44.7	24.5
1970-2	2 860.3	2 020.1	1 087.5	44.4	23.9
1970-3	2 896.6	2 045.3	1 824.7	44.9	23.3
1970-4	2 873.7	2 045.2	1 821.2	42.1	23.1
1971-1	2 942.9	2 073.9	1 849.9	48.8	23.8
1971-2	2 947.4	2 098.0	1 863.5	50.7	23.7
1971-3	2 966.0	2 106.6	1 876.9	54.2	23.8
1971-4	2 980.8	2 121.1	1 904.6	55.7	23.7

1 C.W.J.Granger和P.Newbold，“Spurious Regression in Econometrics，”*Journal of Econometrics*，vol.2,no.2, July 1974，p.111-120。

2 任何时间序列数据都可以被看成是由一个随机的过程和一个具体的数据集所产生的，例如如表14-5的数据，可以看作是一个随机过程的实现。

(续)

季度	GDP	PDI	PCE	利润	红利
1972-1	3 037.3	2 129.7	1 929.3	59.4	25.0
1972-2	3 089.7	2 149.1	1 963.3	60.1	25.5
1972-3	3 125.8	2 193.9	1 989.1	62.8	26.1
1972-4	3 175.3	2 272.0	2 032.1	68.3	26.5
1973-1	3 253.3	2 300.7	2 063.9	79.1	27.0
1973-2	3 267.6	2 315.2	2 062.0	81.2	27.8
1973-3	3 264.3	2 337.9	2 073.7	81.3	28.3
1973-4	3 289.1	2 382.7	2 067.4	85.0	29.4
1974-1	3 259.4	2 334.7	2 050.8	89.0	29.8
1974-2	3 267.7	2 304.5	2 059.0	91.2	30.4
1974-3	3 239.1	2 315.0	2 065.5	97.1	30.9
1974-4	3 226.4	2 313.7	2 039.9	86.8	30.5
1975-1	3 154.0	2 282.5	2 051.8	75.8	30.0
1975-2	3 190.4	2 390.3	2 086.9	81.0	29.7
1975-3	3 249.9	2 354.4	2 114.4	97.8	30.1
1975-4	3 292.5	2 389.4	2 137.0	103.4	30.6
1976-1	3 356.7	2 424.5,	2 179.3	108.4	32.6
1976-2	3 369.2	2 434.9	2 194.7	109.2	35.0
1976-3	3 381.0	2 444.7	2 213.0	110.0	36.6
1976-4	3 416.3	2 459.5	2 242.0	110.3	38.3
1977-1	3 466.4	2 463.0	2 271.3	121.5	39.2
1977-2	3 525.0	2 490.3	2 280.8	129.7	40.0
1977-3	3 574.4	2 541.0	2 302.6	135.1	41.4
1977-4	3 567.2	2 556.2	2 331.6	134.8	42.4
1978-1	3 591.8	2 587.3	2 347.1	137.5	43.5
1978-2	3 707.0	2 631.9	2 394.0	154.0	44.5
1978-3	3 735.6	2 653.2	2 404.5	158.0	46.6
1978-4	3 779.6	2 680.9	2 421.6	167.8	48.9
1979-1	3 780.8	2 699.2	2 437.9	168.2	50.5
1979-2	3 784.3	2 697.6	2 435.4	174.1	51.8
1979-3	3 807.5	2 715.3	2 454.7	178.1	52.7
1979-4	3 814.6	2 728.1	2 465.4	173.4	57.6
1980-1	3 830.8	2 742.9	2 464.6	174.3	57.6
1980-2	3 732.6	2 692.0	2 414.2	144.5	58.7
1980-3	3 733.5	2 722.5	2 440.3	151.0	59.3
1980-4	3 808.5	2 777.0	2 469.2	154.6	60.5
1981-1	3 860.5	2 783.7	2 475.5	159.5	64.0
1981-2	3 844.4	2 776.7	2 476.1	143.7	68.4
1981-3	3 864.5	2 814.1	2 487.4	147.6	71.9
1981-4	3 803.1	2 808.8	2 468.8	140.3	72.4
1982-1	3 756.1	2 795.0	2 484.0	114.4	70.0
1982-2	3 771.1	2 824.8	2 488.9	114.0	68.4
1982-3	3 754.4	2 829.0	2 502.5	114.6	69.2
1982-4	3 759.6	2 832.6	2 539.3	109.9	72.5
1983-1	3 783.3	2 843.6	2 556.5	113.6	77.0
1983-2	3 886.5	2 867.0	2 604.0	133.0	80.5
1983-3	3 944.4	2 903.0	2 639.0	145.7	83.1
1983-4	4 012.1	2 960.6	2 678.2	141.6	84.2
1984-1	4 221.8	3 123.6	2 824.9	125.2	87.2

(续)

季度	GDP	PDI	PCE	利润	红利
1984-2	4 144.0	3 065.9	2 741.1	152.6	82.2
1984-3	4 166.4	3 102.7	2 754.6	141.8	81.7
1984-4	4 194.2	3 118.5	2 784.8	136.3	83.4
1985-1	4 221.8	3 123.6	2 824.9	125.2	87.2
1985-2	4 254.8	3 189.6	2 849.7	124.8	90.8
1985-3	4 309.0	3 156.5	2 893.3	129.8	94.1
1985-4	4 333.5	3 178.7	2 895.3	134.2	97.4
1986-1	4 390.5	3 227.5	2 922.4	109.2	105.1
1986-2	4 387.7	3 281.4	2 947.9	106.0	110.7
1986-3	4 412.6	3 272.6	2 993.7	111.0	112.3
1986-4	4 427.1	3 266.2	3 012.5	119.2	111.0
1987-1	4 460.0	3 295.2	3 011.5	140.2	108.0
1987-2	4 515.3	3 241.7	3 046.8	157.9	105.5
1987-3	4 559.3	3 285.7	3 075.8	169.1	105.1
1987-4	4 625.5	3 335.8	3 074.6	176.0	106.3
1988-1	4 655.3	3 380.1	3 128.2	195.5	109.6
1988-2	4 704.8	3 386.3	3 147.8	207.2	113.3
1988-3	4 779.7	3 443.1	3 170.6	213.4	117.5
1988-4	4 779.7	3 473.9	3 202.9	226.0	121.0
1989-1	4 809.8	3 473.9	3 200.9	221.3	124.6
1989-2	4 832.4	3 450.9	3 208.6	206.2	127.1
1989-3	4 845.6	3 446.9	3 241.1	195.7	129.1
1989-4	4 859.7	3 493.0	3 241.6	203.0	130.7
1990-1	4 880.8	3 531.4	3 258.8	199.1	132.3
1990-2	4 832.4	3 545.3	3 258.6	193.7	132.5
1990-3	4 903.3	3 547.0	3 281.2	196.3	133.8
1990-4	4 855.1	3 529.5	3 251.8	199.0	136.2
1991-1	4 824.0	3 514.8	3 241.1	189.7	137.8
1991-2	4 840.7	3 537.4	3 252.4	182.7	136.7
1991-3	4 862.7	3 539.9	3 271.2	189.6	138.1
1991-4	4 868.0	3 547.5	3 271.1	190.3	138.5

注：GDP —— **国内生产总值，单位：**10**亿美元**(1987 **年美元价**)
PDI —— **个人可支配收入，单位：**10**亿美元**(1987**年美元价**)
PCE —— **个人消费支出，单位：**10**亿年美元**(1987**年美元价**)
利润 —— **公司税后利润，单位：**10**亿美元**(1987**年美元价**)
红利 —— **公司净红利支出，单位：**10**亿美元**(1987**年美元价**)
数据来源：U.S Department of Commerce，Bureau of Economic Analysis，*Business Statistics*,1963-1991,June 1992。

用符号表示，令Y_t**代表随机时间序列，如果它满足以下条件，则该时间序列是平稳的**[1]**：**

均值： $$E(Y_t)=u \tag{14-36}$$

方差： $$E(Y_t-u)^2=\sigma^2 \tag{14-37}$$

协方差： $$\gamma_k=E[(Y_t-u)(Y_{t+k}-u)] \tag{14-38}$$

其中γ_k**表示**Y_t**期与**Y_{t+k}**期的协方差(或者自协方差)，也即相隔**k**期两个**Y**值之间的协方差。如果**$k=0$**，则为**γ_0**，这就是**Y**的方差**$(=\sigma^2)$**；如果**$k=1$**，则**γ_1**是两个相邻**Y**值之间的协方差，这也正是**

1 在时间序列文献中，这样的随机过程被称作是弱平稳随机过程，在许多实际工作中，弱平稳随机过程这一假定是很有用的。在强平稳随机过程中，我们要考虑更高阶的矩，也就是说二次以上的矩。

我们在第12章讨论自相关时遇到的协方差类型。

假设我们将Y的原点从Y_t移到Y_{t+m}(比方说，在这个例子中从1970-1移到1974-1)。若Y_t是平稳的，则Y_{t+m}的均值、方差和自协方差应该和Y_t相同。简言之，如果一个时间序列是平稳的，其均值、方差和自协方差都是相同的(无论在什么时期去度量)。

如果一个时间序列不是上述定义的平稳时间序列，则称为非平稳时间序列(需要记住的是，我们讨论的仅仅是弱稳定)。

我们来看图14-3的PCE和PDI时间序列图，我们似乎觉得这两个时间序列并不是平稳的。如果情形果真是这样，则在回归(14-35)中，我们是以一个非平稳时间序列来拟合另一个非平稳时间序列，这会产生伪回归现象。

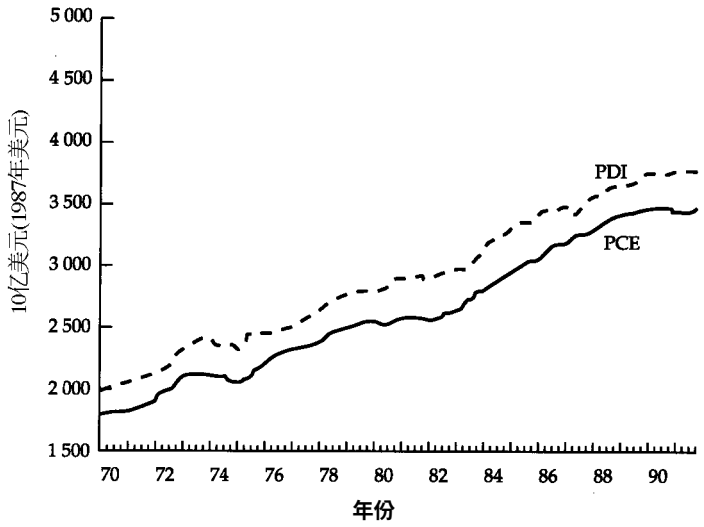


图14-3　1970~1979年美国的PDI和PCE(季度数据)

现在的问题是如何来验证我们的“印象”——PCE和PDI这两个时间序列确实是非平稳的？

14.6.1　非平稳检验：单元根检验

非平稳检验的方法有多种。这里，我们介绍一种比较新的检验方法，单元根检验(unit root test)。其检验步骤如下(具体的技术细节从略)：[1]设Y_t代表利率这一随机时间序列(例如PCE)

(1) 估计回归方程：

$$\Delta Y_t = A_1 + A_2 t + A_3 Y_{t-1} + u_t \tag{14-39}$$

其中Δ是一阶差分符号，在第12章已经遇到过，其中t是趋势变量，取值为1、2等等(在我们这个例子到88)，Y_{t-1}为变量Y的一期滞后值。[2]

(2) 零假设为Y_{t-1}的系数A_3为零，这等价于时间序列是非平稳的。我们称这个假设为单元根假设。[3]

1　详细的讨论参见Damodar N. Gujarati，*Basic Econometrics*，3d ed., McGraw-Hill，New York，1995，Chap.21。

2　回归方程(14-39)也可以不包括截距和趋势变量，尽管通常都包括。

3　为了直观地说明为什么使用单元根这一术语，我们作如下说明：$Y_t = A_1 + A_2 t + CY_{t-1} + u_t$。现在在方程的两边同时减去$Y_{t-1}$得到$(Y_t - Y_{t-1}) = A_1 + A_2 t + CY_{t-1} - Y_{t-1}$，也就是$\Delta Y_t = A_1 + A_2 t + (C-1)Y_{t-1} = A_1 + A_2 t + A_3 Y_{t-1}$，其中$A_3 = C - 1$。这样，如果$C$等于1，则回归方程(14-39)中的$A_3$将为零。这也就是单元根这一术语的由来。

(3) 为了检验A_3的估计值a_3为零，通常我们都会使用很熟悉的 t检验。但不幸的是，这里，我们不能这么做，因为严格来说，t检验只有当时间序列是平稳的才有效。但是可以使用另一种检验方法，τ检验，它的临界值由这一方法的提出者根据蒙特卡罗模拟给出。在文献中， τ检验也被称作是Diskey－Fuller(DF)检验。[1] 若在应用中，估计得到的A_3的τ值(绝对值)大于临界值的D-Fτ值，将拒绝单元根假设，也就是说，我们判定该时间序列是稳定的。另一方面，若计算的τ值(绝对值)小于临界的τ值，则不能拒绝单元根假设。在这种情况下，该时间序列是非平稳的。

下面，我们将单元根检验用于表 14-5给出的PCE和PDI时间序列。与方程 (14-39)相对应，有：

$$\Delta PCE_t = 93.392 + 0.798\ 7t - 0.044\ 5PCE_{t-1} \tag{14-40}$$

$$t(=\tau) = (1.678)\quad (1.360)\quad (-1.376)\qquad R^2 = 0.022\ 1$$

$$\Delta PDI_t = 326.63 + 2.875\ 2t - 0.1569PDI_{t-1} \tag{14-41}$$

$$t(=\tau) = (2.755)\quad (2.531)\quad (-2.588)\qquad R^2 = 0.075\ 7$$

我们现在关注的是滞后PCE和PDI的τ值。MacKinnon计算得到在1%，5%，10%显著水平下，临界的DF值分别为－4.067，－3.460 0和－3.244 7。[2] 滞后的PCE和PDI变量的τ值(绝对值)比上述任何一个τ值都小，由此得出结论： PCE和PDI时间序列是非平稳的 (也就是说有一个单元根)。因此，方程(14-35)给出的OLS回归结果可能是虚假的 (也即没有意义的)。顺便指出：如果用通常的t检验，则滞后 PDI的t值是统计显著的。但是根据 τ检验(当时间序列是非动态的)，这一结论是错误的。

14.6.2 协整时间序列

回归(14-35)可能是伪回归这一结论或许表明所有类似方程 (14-35)的时间序列回归都是虚假的。如果事实的确如此，那么在作时间序列回归分析时要特别谨慎。但没有理由绝望。即使时间序列PCE和PDI不是平稳的，这两个变量之间仍然可能存在一种 (长期)稳定的或均衡的关系。如果是这样的话，我们就说该时间序列是协整的(cointegration)。[3] 但是，如何确定这一点呢？我们来看下面的例子。

从PCE-PDI回归方程(14-35)中，得到残差，e_t：

$$e_t = PCE_t + 171.441\ 2 - 0.967\ 2PDI_t \tag{14-42}$$

把e_t看作一个时间序列，利用单元根检验 [参见方程(14-39)]，得到如下结果(注意：在这一回归中无需引入截距和趋势变量 (为什么？))

$$\Delta\hat{e}_t = -0.275\ 3e_{t-1} \tag{14-43}$$

$$t(=\tau) = (-3.779)\qquad r^2 = 0.142\ 2$$

Engle和Grangeer计算的临界τ值为－2.589 9 (5%)，－1.949 3(5%)和－1.617 7(10%)。[4] 从(14-13)得到的τ的绝对值为3.779，远超过这些临界的τ值。由此得出结论：时间序列 e_t是平稳的。因而我们可以说，尽管时间序列 PCE和PDI各自并不是平稳的，但它们的线性组合 [如方程(14-42)所示]却是平稳的。也就是说，这两个时间序列是协整的，换言之，这两个变量之间存在着长期

1 D. A. Dickey和W. A. Fuller,“Distribution of Estimator for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association*，vol.74，June 1979，p.427-431。

2 J. G. MacKinnon，“Critical Values of Cointegration Tests,” in R. F. Engle , C. W. J. Granger, *Long-Run Economic Relationships*: *Readings in Cointegration*, ，Oxford University Press, New York，1991，Chap.13。现在，计算机软件包例如EVIEWS，可以计算τ值。

3 有关协整的文献非常多，也比较专深。我们这里的讨论只是起一个抛砖引玉的作用。

4 R.F.英格(R. F. Engle)和C.W.J.格朗(C. W. J. Granger)，同前。

的或均衡的关系。这是一个令人欣慰的结果，因为它意味着回归结果 (14-35)是真实的而不是虚假的。

总之，在处理时间序列数据时，我们必须确保要么每个时间序列是平稳的，要么它们是协整的。否则就可能陷入伪(无意义的)回归分析。

下面，我们考虑另一类非平稳时间序列 —— 随机游走模型，已经证实：在金融、投资和国际贸易领域，随机游走模型非常有用。

14.6.3 随机游走模型

金融时间序列，例如 S&P500股票指数、道 -琼斯指数、外汇汇率等等通常被认为是服从“ 随机游走 ”的，这主要是就这种意义而言的，即根据变量的今天的值并不能使我们预测这些变量的明日值。因此，知道今天股票的价格 (比方说是康柏计算机或者 IBM)，却很难知道明天它会是多少。也就是说，股票价格行为本质上是随机的 —— 今天的价格等于昨天的价格加上一个随机冲击。[1]

为了剖析一个随机游走模型，考虑如下简单模型：

$$Y_t=Y_{t-1}+u_t \tag{14-44}$$

其中u_t是均值为零方差固定为σ^2的随机误差项，让我们假设我们在时间 0开始值为Y_0。现在我们可以写成：

$$Y_{t-1}=Y_{t-2}+u_{t-1} \tag{14-45}$$

利用递回关系(14-44)重复方程(14-45)所做，我们可以写成：

$$Y_t=Y_0+\sum u_t \tag{14-46}$$

其中加总是从t=1到t=T，T是观察总次数。现在可以很容易地核实

$$E\ (Y_t)=Y_0 \tag{14-47}$$

因为每个u_t的期望值都为零，而且也很容易核实

$$\text{var}\ (Y_t)=\text{var}\ (u_t+u_{t-1}+\cdots+u)=T\sigma^2 \tag{14-48}$$

其中我们利用了这一事实，即u是随机的，每个都有同样的方差σ^2。

正当方程(14-48)所示，Y_t的方差不仅不是固定的，而且是随着 T的增加而不断地增加的。因此，依据我们前面所给出的静态时间序列的定义，方程 (14-44)中所给的(随机游走)变量Y_t是非静态时间序列(这里，非静态是指在方差中)。但是注意方程(14-44)中所给的模型的一个有意思的特征。如果你将它写成：

$$\Delta Y_t=(Y_t-Y_{t-1})=u_t \tag{14-49}$$

如以前一样，其中 Δ是操作数的一阶差分，我们看到， Y的一阶差分是静态的，因为 $E\ (\Delta Y_t)=E\ (u_t)=0$并且有$\text{var}\ (\Delta Y_t)=\text{var}(u_t)=\sigma^2$。因此，如果方程(14-44)中的$Y$表示，比方说是股票价格，则这些价格可能是非静态的，但是它们的一阶差分却纯粹是随机的。

我们可以将随机游走模型 (14-44)调整为：

$$Y_t=d+Y_{t-1}+u_t \tag{14-50}$$

其中d为一常数。这就是带漂移项的随机游走模型(Random walk model with drift)或带常数项的随机游走模型，d为漂移参数。

留给读者证明：

$$E\ (Y_t)=Y_0+Td \tag{14-51}$$

---

1 随机游走经常与醉鬼游走相比较。在离开酒馆时，醉鬼在时间 t移动一个随机的距离u_t，如果他或者她无限地走下去，他或者她最终会离酒馆越来越远。

$$\text{var}(Y_t)=T\sigma^2 \tag{14-52}$$

也就是说，对于带漂移项的随机游走模型来说，均值和方差都随着时间不断增加。再一次，我们又有了一个非平稳的随机变量——无论是均值还是方差都是非平稳的。如果 d是正的，从模型(14-50)中看到，Y的均值会随时间不断地增加。如果 d是负的，则 Y的均值会不断地减少。在任何一种情形下，Y的方差都是随时间不断地增加。一个随机变量，如果它的均值和方差是与时间有关的，则我们说它服从随机趋势(stochastic trend)。这与我们在第8章中所讨论的线性趋势模型是[参见方程(8-23)]相对照的，在那里我们假设变量 Y服从确定性趋势(deterministic trend)。

如果用随机游走模型进行预测，我们获得如图 14-4所示的图形。图14-4a表明的是不带漂移项的随机游走模型，而图 14-4a所示的是带漂移项的随机游走模型。我们能够看到，在图 14-4a中平均预测值在整个未来都保持着 Y_T时的水平，但是由于方差的增加，围绕均值的置信区间则在不断地加大。在图 14-4b中，假设漂移项参数是正的，则平均预测值随时间是增加的。预测误差也在增加。

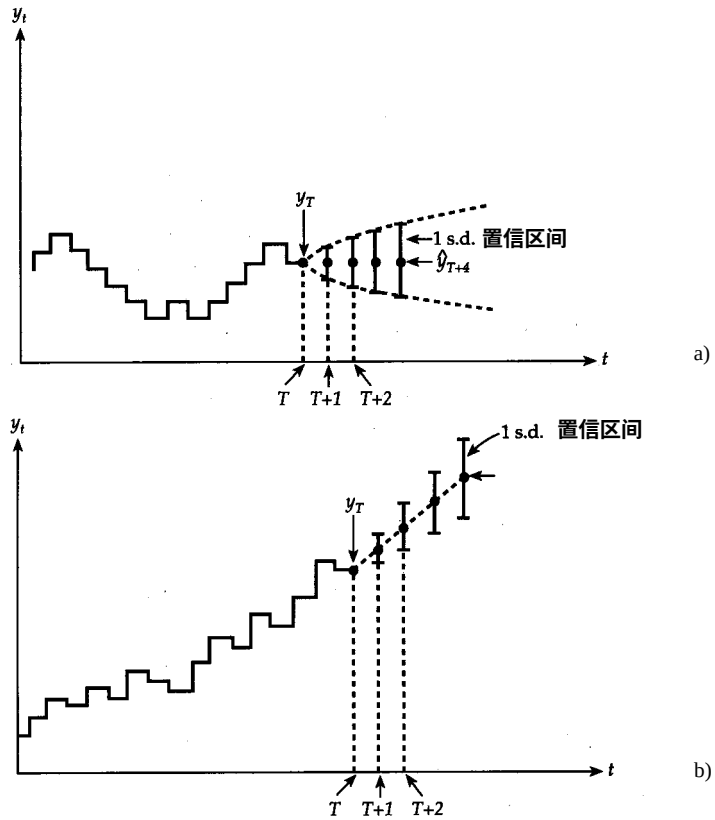


图14-4 利用随机游走模型进行预测

资料来源：Adapted from Robert S. Pindyck and Daniel L. Rubinfeld, Econometric Models and Economic Forecasts 4th ed., McGraw-Hill, New York, 1998, pp. 491-492.

总之，我们一直是想提醒读者在建立涉及时间序列数据的模型时需要谨慎。如果被解释变量Y和解释变量X是非平稳的，则较高的 R^2和较高的t值可能会使你误以为你已经找出了两者之间的一种有意思的关系。事实上，较高的 R^2可能仅仅反映了这一事实，即两个变量拥有相同的趋势，因此它们之间没有任何真正的关系。这就是伪回归现象。根据Granger和Newbold的理解，

判断伪回归是的一个依据是：在涉及时间序列数据的回归中， R^2值远大于杜宾-瓦尔森d值。因此，要注意这一点。

14.7 小结

在这一章里，我们讨论了四个很有实践意义的专题。第一个专题是有限最小二乘法 (RLS)。如果经济理论表明回归模型的一个或者多个参数必须满足一个或者多个线性约束， RLS方法把这类约束直接加入到估计过程之中。我们将 RLS回归结果与ULS回归结果(或者说普通的OLS回归结果)相比较，包括比较有限回归和无限制回归的 R^2值。在约束条件是有效的这一零假设下，通过方程(14-8)说明了F检验如何帮助我们决定是否拒绝零假设。若不拒绝零假设，则说明理论上所表明的约束条件是有效的；但若拒绝零假设，则说明理论上所表明的约束条件是无效的。在这种情况下，我们可以使用无约束的或者标准的 OLS法。

我们讨论的第二个专题是动态建模，就是把时间或者滞后因素引入模型。在这类模型中，被解释变量的当前值依赖于一个或者多个解释变量的滞后值。这种依赖可能是是由于心理、技术或者制度等等原因。通常称这类模型为分布滞后模型。尽管说，引入一个或者多个滞后解释变量并不违背标准的CLRM假定，但是我们不提倡用 OLS方法来估计这类模型，原因在于多重共线性问题以及每增加一个估计系数就意味着失去一个自由度。因此，对这类模型的估计通常是通过给模型的参数附加一些约束条件 (例如滞后系数值从第一项起由前往后递减)来实现。这就是夸克模型，适应性预期模型，部分或存量调整模型所采纳的方法。这些模型的一个显著特征是它们用一个被解释变量的滞后值代替了所有解释变量的滞后值。由于被解释变量滞后值在模型中是以解释变量的形式存在，因此，这类模型称为自回归模型。尽管自回归模型在估计分布滞后系数方面取得了成功，但它并不是没有任何统计问题的。特别地，我们必须小心误差项可能存在自相关，因为在存在自相关以及在应变量滞后值作为解释变量的情形下， OLS估计量是有偏的和不一致的。

在讨论动态模型时，我们指出了它们是如何帮助我们估计解释变量对被解释变量的短期和长期影响。

接下来我们讨论的了取值为 0、1的虚拟应变量。尽管这类模型可用 OLS法加以估计(这种情况下，称作线性概率模型)，但我们建议最好不用，因为这类模型估计得到的概率有时可能为负或大于1。因此，通常用分对数模型或概率模型来估计。本章，我们通过几个具体的例子说明了分对数模型。许多优秀统计软件的出现，使得估计分对数模型和概率模型不再是一件神秘和令人生畏的工作。

我们所讨论的最后一个专题是伪回归现象 (无意义回归)。当我们用一个或者多个非平稳随机变量来回归另一个非平稳随机会变量时，就产生了伪回归。一个时间序列被称作是 (弱)稳定的，如果它在不同时滞里的均值、方差和协方差都与时间无关。为了确定一个时间序列是否为平稳的，我们可以使用单元根检验。如果单元根检验 (或者其他检验)表明我们所关心的时间序列是平稳的，则根据这样的时间序列得到的回归结果就不是虚假的。

我们还引入了协整的概念。两个或者多个时间序列被称作是协整的，如果尽管单独来看两个时间序列是非平稳的、但它们之间却存在着稳定的、长期的关系。如果是这样的话，则根据这样的时间序列得到的回归结果就不是虚假的。

最后我们介绍了带或者不带漂移项的随机游走模型。许多金融时间序列都认为是服从随机游走的，也就是说，它们是非平稳的，这或者反映在它们的均值上，或者反映在它们的方差上，或者两者都有所反映。具有这些特征的变量被称作是服从随机趋势。股票价格是随机游走的一

个典型例子。知道今天的股票价格很难预知明天的股票价格。关于明天价格的最佳的猜测是今天的价格加上或者减去一个随机误差项 (或者通常所说的冲击)。如果我们能够非常精确地预测明天的股票价格，那我们就都将成为百万富翁了。

习题

14.1 解释下列术语：

(a)有限最小二乘法(RLS) (b) 动态模型 (c)分布滞后模型 (d)自回归模型

14.2 通常的情形是$R^{*2} \leq R^2$；也就是说，从RLS中获得的R^2小于或者至多等于从ULS中所获得的R^2。你能够对此提供一个直观的理由吗？

14.3 如果理论上所附加的限制事实上是无效的， RLS估计量将会比无限制的OLS估计量更加有效率。你能够直观地解释这一点吗？

14.4 考虑模型：

模型A：$Y_i=B_1+B_2X_i+u_i$

模型B：$Y_i=AX_i+v_i$

(a) 两个模型之间的区别是什么？ (b) 你如何利用RLS来确定在给定情形下哪个是恰当的模型？ (c) 假设模型A是正确的模型但是你却估计了模型B。这其中涉及哪种类型的设定偏差？你如何利用RLS来诊断这一偏差？

14.5 考虑以下模型：

$$Y_i=B_1+B_2X_{2i}+B_3X_{3i}+B_4X_{4i}+B_5X_{5i}+u_i$$

假设你想附加限制条件$B_2=0$和$B_4+B_5=1$，你如何建立RLS模型？你又如何检验这些限制的有效性？

14.6 被解释变量对于一个或者多个解释变量反应滞后的原因是什么？给出一些分布滞后模型的例子。

14.7 在一个分布滞后模型里继续地也就是说如果一个增加的滞后项的t值在统计上是显著的就加入每一个后继变量。换言之，只要增加的滞后项的t值根据t检验在统计上是显著的，则继续加入一个滞后项，以此来确定滞后项的数量。这种策略有什么错误？

14.8 “由于在分布滞后模型中的连续滞后项之间很有可能是共线性的，在这样的模型中我们不应该去担心任何个别滞后项系数的统计显著性，而应该考虑滞后系数加总作为一个整体的统计显著性，”就此论断作出评论。

14.9 “尽管分对数模型和概率模型可能会优于LPM模型，然而在实践中，正如‘简单优于复杂’，我们应该选择LPM原因在于其简单性。”你同意吗？为什么？

14.10 判断正误：分对数的值越大，则特定事件发生的概率越大。

14.11 对应于回归(14-3)，以下是1958年到1972年期间台湾经济农业部门的回归结果：

$$\ln \hat{Y}_t = -3.338\ 4+1.498\ 8\ln X_{2i}+0.489\ 9\ \ln X_{3i}$$

$$t=(-1.362\ 9)\quad (0.539\ 8)\quad (0.102\ 0)\qquad R^2=0.889\ 0$$

附加限制条件为：$B_2+B_3=1$，也就是说，规模收益不变，可以得出如下回归：

$$\ln \frac{\hat{Y}_t}{X_{2t}} = 1.708\ 0-0.613\ 0\ \ln \frac{X_{3t}}{X_{2t}}$$

$$t=(4.108\ 2)\quad (6.570\ 2)\qquad R^{*2}=0.848\ 9$$

(a) 解释这两个回归。 (b) 给定上述回归结果，你能说在台湾农业部门中是否存在着固定规模收益吗？给出计算步骤。 (c) 农业部门的结果与课文中所给的制造业部门的结果有何

不同？

14.12 表14-6给出了1970年到1987年期间美国的个人消费支出 (PCE)和个人可支配收入(PDI)数据，所有数字的单位都是10亿美元(1982年的美元价)。

表 14-6

年份	PCE	PDI	年份	PCE	PDI	年份	PCE	PDI
1970	149 2.0	166 8.1	1976	180 3.9	200 1.0	1982	205 0.7	226 1.5
1971	153 8.8	172 8.4	1977	188 3.8	206 6.6	1983	214 6.0	233 1.9
1972	162 1.9	179 7.4	1978	196 1.0	216 7.4	1984	224 9.3	246 9.8
1973	168 9.6	191 6.3	1979	200 4.4	221 2.6	1985	235 4.8	254 2.8
1974	167 4.0	189 6.6	1980	200 0.4	221 4.3	1986	245 5.2	264 0.9
1975	171 1.9	193 1.7	1981	204 2.2	224 8.6	1987	252 1.0	268 6.3

数据来源：《总统经济报告，1989》，PCE从第325页的表B-15中而来，PDI则从第339页的表B-27中来。

估计下列模型：

$$PCE_t=A_1+A_2PDI_t+u_t$$

$$PCE_t=B_1+B_2PDI_t+B_3PCE_{t-1}+v_t$$

(a) 解释这两个回归模型的结果。

(b) 短期和长期边际消费倾向 (MPC)是多少？

14.13 利用题14.12中的数据，考虑以下模型：

$$\ln PCE_t=A_1+A_2\ln PDI_t+u_t$$

$$\ln PCE_t=B_1+B_2\ln PDI_t+B_3\ln PCE_{t-1}+v_t$$

其中ln=自然对数。

(a) 阐述这些回归。

(b) PCE对于PDI的短期和长期弹性是多少？

14.14 为了估量设备利用对于通货膨胀的影响，托马斯 A.吉廷斯[1] 根据1971年到1988年的美国数据获得如下回归：

$$\hat{Y}_t=-30.12+0.140\ 8X_t+0.236\ 0X_{t-1}$$

$$t=(-6.27)\quad(2.60)\quad(4.26)\qquad R^2=0.727$$

其中，Y=GNP隐含折算数，%(通货膨胀率的一种测度)

X_t=制造业设备利用率，%

X_{t-1}=滞后一年的设备利用率

(a) 阐述上述回归。先验地，为什么在通货膨胀和设备利用之间存在着正相关。 (b) 设备利用对于通货膨胀的短期影响是多少？长期影响又是多少？ (c) 每个斜率系数是统计显著的吗？ (d) 你是否会拒绝零假设：两个斜率系数同时为零？你将利用何种检验？ (e) 如果有数据进行如下回归分析

$$Y_t=B_1+B_2X_t+B_3Y_{t-1}+u_t$$

你如何找出设备利用对于通货膨胀的短期和长期影响？你能够将这一模型的结果与吉廷斯所给出的结果作一比较吗？

14.15 表14-7给出了喷射不同浓度的洛特纳 (rotenone，一种农药)对于50批菊花蚜虫的影

1 Thomas A. Gittings，“Capacity Utilization and Inflation”，*Economic Perspectives*，Federal Reserve Bank of Chicago，May/June 1989，p.2-9。

响的数据。

表 14-7

浓度(毫克／升) X	$\log(X)$	总数 N_i	死亡 n_i	$p_i = \frac{n_i}{N_i}$
2.6	0.415 0	50	6	0.120
3.8	0.579 7	48	16	0.333
5.1	0.707 6	46	24	0.522
7.7	0.886 5	49	42	0.857
10.2	1.008 6	50	44	0.880

注：对数是普通对数，即以10为底的对数。

数据来源：D. J. Finney，*Probit Analysis*，Cambridge University，London，1964。

建立一个合适的模型表示死亡率对 $\log X$ 的函数，并对结果作评论。

14.16 考虑如下模型，这一模型也被称作是逻辑斯蒂 (概率)分布，其中 P_i 是拥有一套住房的概率，X_i 为收入。

$$P_i = \frac{1}{1+\exp\{-B_1 - B_2 X_i\}}$$

其中exp{}表示 e 的{}次幂。通过简单的代数变换，我们可以得到：

$$\ln\left(\frac{P_i}{1-P_i}\right) = B_1 + B_2 X_i$$

这就是课文中所讨论过的分对数模型。这一名称来自于上述逻辑斯蒂分布。

14.17 根据表14-2中所给的数据和回归(14-26)，计算每个学生获得A的概率。

14.18 从回归(14-33)中，计算表14-3中每个收入水平上拥有住房的概率。

14.19 根据一个20对夫妇的样本，巴巴拉·邦德·杰克逊 (Barbara Bund Jackson) [1]获得了如下回归：

$$\ln\frac{P_i}{1-P_i} = 0.945\ 6 + 0.363\ 8\text{收入} - 1.107\text{保姆}$$

(注：作者并没有给出标准差。)

其中，

P=利用餐馆的概率，＝1，如果去餐馆就餐的话，否则等于0。

收入＝以千美元计算的收入

保姆＝1，如果雇用了保姆的话，否则等于0

在20对夫妇中，有11对夫妇经常去餐馆就餐，6对经常雇用保姆。收入范围从低的17 000美元到高的44 000美元。

(a) 解释上述分对数回归。 (b) 求收入为44 000的夫妇需要保姆的分对数值。 (c)对于同样的夫妇，求出他们去餐馆就餐的概率。

14.20 参考表14-5中所给的数据。

(a) 画出利润和红利的散点图，并直观地检验一下这两个时间序列是否是稳定的。 (b) 应用单元根检验分别检验这两个时间序列是否是平稳的。 (c) 如果利润和红利时间序列并不是平稳的，而如果你以利润来回归红利，那么回归的结果会是虚假的吗？为什么？你是如何判定的？说明必要的计算。 (d) 取这两个时间序列的一阶差分并确定这个一阶差分时间序列是

1 参见Barbara Bund Jackson，Multivariate Data Analysis: An Introduction，Irwin，1983，p.92。

否是平稳的。

14.21　蒙特卡罗试验。考虑以下随机游走模型：

$$Y_t=Y_{t-1}+u_t$$

假设$Y_0=0$并且u_t是独立的正态分布，其均值为0，方差为9。产生100个u_t值，并利用这些值产生100个Y_t值。对所获得的Y值作散点图。根据散点图你得出什么结论？

14.22　蒙特卡罗试验。现在假设有以下模型：

$$Y_t=4+Y_{t-1}+u_t$$

其中$Y_0=0$，并且u_t与题14.21中所述的u_t是一样的。重复题14.21中的过程。这一试验与前一试验有何区别？

第四部分

联立方程模型简介

本书前三部分主要讨论了单方程回归模型，这些模型在经济和商业领域的实际工作中得到广泛的运用。我们知道，在单方程模型中，一个变量(应变量Y)可以表示为一个或多个变量(解释变量，X)的线性函数。在这类模型中的一个暗含假定是，如果Y与X之间存在因果关系，则这种关系是单向的，即解释变量是原因，被解释变量(应变量)是结果。

但是，有些时候经济变量之间的影响却是双向的。也就是说，一个经济变量影响另一个经济变量(或多个变量)，反过来，又受其他经济变量的影响。例如，在货币(M)对利率(r)的单方程回归模型中，暗含地假定了在利率固定不变的条件下(由联邦储备系统规定)，不同利率水平下的货币需求量的变化。但是，如果利率依赖于货币的需求量，情况又会怎样呢？在这种情况下，条件回归分析就不适用了，因为现在M依赖于r，而r也依赖于M。这就需要我们考虑**联立方程模型**(simultaneous equation models)——模型中包含不止一个的回归方程，也即变量之间的影响是相互的。

在第四部分中，我们简单介绍了联立方程模型的基本知识，有关联立方程模型更详细的讨论可参阅本书最后列出的部分参考文献。

第15章我们首先讨论了联立的含义并指出为什么在前面章节中广泛使用的最小二乘法在估计联立方程模型的参数是不适用的。我们还介绍了估计联立方程模型参数的两种不同方法：间接最小二乘法(ILS)和两阶段最小二乘法(2SLS)，并分别指出了这两种方法的优缺点。

本章我们还讨论了模型识别问题。这个问题的难点在于：在一个联立方程模型中，有两个或更多个方程看似相同以致于我们无法辨别哪一个是我们想要估计的。比如，在包括供给方程和需求方程的模型中，如果我们仅有商品价格及其需求量的数据，则很难辨别哪一个是需求函数，哪一个是供给函数。为了确认，我们需要一些额外的信息。本章我们将利用阶条件来辨别模型是否是可识别的。

本章我们仍将用一些数字例子和具体的经济实例来说明在本章出现的一些基本概念和技术方法。

第15章

联立方程模型

到目前为止，我们所遇到的模型都是单方程模型，即一个应变量 *Y* 是一个或多个解释量 *X* 的函数。相关的经济理论决定了为什么把 *Y* 看做应变量，而把 *X* 看做解释变量。换句话说，在单方程回归模型中，由 *X* 导致了 *Y*(如果这种因果关系存在的话)。例如7.4节的抵押贷款一例，经济理论表明了收入和抵押成本决定了对抵押贷款的需求。

但是，有些情况 *Y* 与 *X* 之间并不能维持这种单向关系。很可能 *X* 不但影响 *Y*，而且 *Y* 也影响 *X*。如果这样的话， *Y* 与 *X* 之间就存在着双向关系(bilateral)或反馈关系(feedback)。显然，在这种情况下，单方程模型就无法准确地描述这种双向关系(也即会导致结果有偏差)。因此，我们需要多个回归方程。包括多个方程，并且变量之间存在反馈关系的回归模型称为联立方程回归模型。本章主要讨论联立方程模型的基本性质。对于联立方程更详细的讨论，读者可参阅有关参考书。[1]

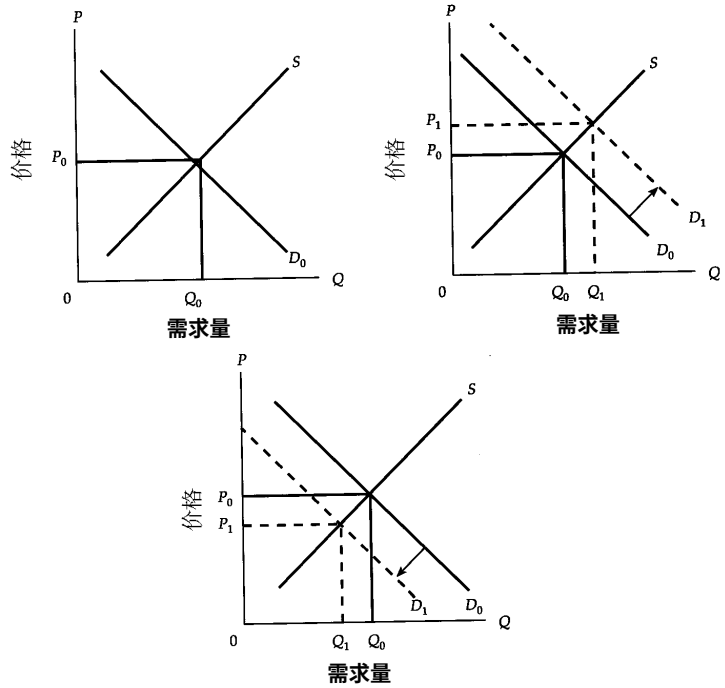


图15-1　价格和需求量的相互影响

1　参见，Damodar N.Gujarati, *Basic Econometrics*, 3d ed., McGraw-Hill, New York, 1995, Chaps. 18-20。

15.1 联立方程模型的性质

我们用几个具体的经济实例来说明联立方程模型的概念。

例15.1 凯恩斯收入决定模型

初学经济学的同学都会接触到简单的凯恩斯收入决定模型。按照标准宏观经济学教科书的规定，用C表示消费(支出)，Y表示收入，I表示投资(支出)，S表示储蓄。简单凯恩斯收入决定模型包括以下两个方程：

消费函数：$C_t=B_1+B_2Y_t+u_t$ (15-1)

收入恒等式：$Y_t=C_t+I_t$ (15-2)

其中，t是时间下标，u是随机误差项，并且，$I_t=S_t$。

简单的凯恩斯模型假定一个封闭的经济（即没有对外贸易），而且没有政府的支出。[回顾一下：通常的收入恒等式写为：$Y_t=C_t+I_t+G_t+NX_t$，其中，G_t表示政府支出，NX_t表示净出口(出口-进口)]。模型还假定了投资支出是由外生决定的，比如说民间投资。

消费函数表明消费支出与收入线性相关；函数中加入随机误差项反映出在经验分析中，两者之间的一种近似关系。国民收入恒等式表明总收入等于消费支出与投资支出之和，而后者又等于储蓄。我们知道，消费函数中的斜率B_2代表了边际消费倾向(MPC)，即由额外一美元收入导致的额外消费支出的数量。凯恩斯假定MPC为正且小于1，这个假设是合理的，因为人们可能会将增加的收入部分地用于储蓄。

现在我们就能够看出消费支出与收入之间的反馈或联立关系。根据式(15-1)可知，收入影响消费支出，但根据式(15-2)，消费又是收入的组成部分。因此，消费和收入是相互影响的。我们分析的目的是想知道消费和收入是如何同时决定的。因而，消费和收入是联合相关变量。用联立方程建模的语言，称联合相关变量为内生变量(endogenous)。在简单凯恩斯模型中，投资I不是内生变量，因为其值是独立决定的，所以称之为外生变量(exogenous)或预定变量(predetermined)。在更复杂的凯恩斯模型中，投资也可成为内生变量。

总之，内生变量是“被研究系统的内在组成部分，并且是在系统内决定的”。换句话说，“内生变量是由因果系统内其他变量所产生的变量”。外生变量是“系统外决定的变量，外生变量不受因果系统的影响。”[1]

方程(15-1)、(15-2)表示了一个包含两个内生变量C和Y的双方程模型。如果有更多个内生变量，那么将有更多个方程，也即有多少个内生变量就有多少个方程。系统中有些方程是结构(structural)方程或行为(behavloral)方程，有些则是恒等式(identity)。在简单的凯恩斯模型中，方程(15-1)就是结构方程或行为方程，因为它描述了经济中某一部分的结构或行为（这里是消费部分)。结构方程中的系数(或参数)，例如B_1、B_2，称为结构系数(structural coefficients)。方程(15-2)是一个收入恒等式(定义式)：总收入等于总消费加上总投资。

例15.2 需求和供给模型

学过经济学的同学都知道，商品的价格P与需求量Q是由需求曲线和价格曲线的交点决定的。这里我们简单地假定需求曲线和供给曲线为直线 (即需求量、供给

1 W.Paul Vogt, *Dictionary of Statistics and Methodology: A Nontechnical Guide for th Social Sciences*, Sage Publications, California, 1993, pp.81, 85。

(续)

量分别与价格线性相关)，加上随机项u_1,u_2，得到下面的需求和供给函数：

需求函数： $Q_t^d = A_1 + A_2P_t + \mu_t$ (15-3)

供给函数： $Q_t^s = B_1 + B_2P_t + \mu_{2t}$ (15-4)

均衡条件： $Q_t^d = Q_t^s$ (15-5)

其中，Q^d表示需求量，Q^s表示供给量，t表示时间。

根据经济理论，预期A_2为负(向下倾斜的需求曲线)，B_2为正(向上倾斜的供给曲线)。方程(15-3)、(15-4)都是结构方程，前者描述了消费者行为，后者描述了供给者行为。A, B是结构系数。

现在就不难理解为什么价格P和需求量Q之间存在联立或双向关系了。举个例子，由于影响需求的其他变量发生变化(例如收入、财富和偏好)，从而使得u_{1t}[式(15-3)]发生变化，如果u_{1t}为正，则需求曲线向上移动，如果u_{1t}为负，则需求曲线向下移动。图15-1表明，需求曲线的移动将导致需求量和价格发生变化。类似地，u_{2t}(由于罢工、天气、飓风)的变化将使供给曲线移动，因而又会影响到价格和需求量。因此，这两个变量之间的关系是双向的；因而，价格和需求变量即是联合相关变量或内生变量。这就是我们所说的联立问题。

15.2 联立方程的偏误：普通最小二乘估计量的不一致性

为什么存在联立问题呢？我们还是回到例15.1的简单凯恩斯收入决定模型。假设暂时忽略消费支出和收入之间的联立性，通过普通最小二乘法仅仅估计消费函数，利用式(5-17)，得到：

$$b_2 = \frac{\sum (C_t - \bar{C})(Y_t - \bar{Y})}{\sum (Y_t - \bar{Y})^2} \tag{15-6}$$

在第6章我们讨论过，如果满足古典线性回归模型的基本假定，那么利用普通最小二乘法得到的估计量是最优线性无偏估计量(BLUE)。[1] 式(15-6)中的b_2是真实的边际消费倾向B_2的最优线性无偏估计量吗？可以证明：一般地，当存在联立问题时，普通最小二乘估计量不是最优线性无偏估计量。在我们这个例子中，b_2就不是B_2的最优线性无偏估计量。而且，b_2是B_2的有偏估计量；平均而言，它高估或低估了真实的B_2。详细的证明参见附录15A。不过从直观上很容易看出为什么b_2不是最优线性无偏估计量。

在第6章中我们讨论过古典线性回归模型的假定之一：随机误差项u与解释变量不相关。因而，如果我们要用普通最小二乘法估计消费函数(15-1)中的各个参数，则要求在凯恩斯消费函数中，Y(收入)与误差项u_t不相关。但在这里却不是的：

$$\begin{aligned} Y_t &= C_t + I_t \\ &= (B_0 + B_1Y_t + u_t) + I \qquad \text{(15-1)式的} C_t \text{带入} \\ &= B_0 + B_1Y_t + u_t + I_t \end{aligned} \tag{15-2}$$

把B_1Y_t项移到等式左边，合并同类项，整理得：

$$Y_t = \frac{B_0}{1 - B_1} + \frac{1}{1 - B_1} I_t + \frac{1}{1 - B_1} \mu_t \tag{15-7}$$

1 这是6.3节中讨论的高斯-马尔柯夫定理的核心。该定理提供了使用最小二乘法的理论依据。

注意观察这个等式，你会发现一条用趣的性质：国民收入Y不但取决于投资I，而且还取决于随机误差项u！我们知道随机误差项u代表了未明确包括在模型之中的各影响因素。假定消费者信心(用密芝安大学建立的消费者信心指数度量)是其中的一个影响因素。若消费者由于股票市场的繁荣(比如美国1996年，1997年)而对经济持乐观态度，则消费者将增加消费支出，根据收入恒等式，它将影响到收入Y。根据消费函数，收入的增加又导致新一轮消费的增加，如此继续。这个过程的最终结果是什么呢？熟悉宏观经济学的学生将会认识到，最终结果取决于乘数$\frac{1}{1-B_2}$的大小。举个例子，如果MPC (B_2)是0.8(也就是说，每增加一美元将有80美分用于消费)，则乘数为5。

这里有一点需要指出：在式 (15-1)中，Y与u是相关的，因此，我们不能用普通最小二乘法估计消费函数中的参数。如果我们坚持要用，那么估计量将是有偏的，而且甚至是不一致的(证明见附录15A)。一般地，若估计量即使随着样本容量的无限增大，也不接近真实的参数值，那么就称这个估计量是非一致估计量(inconsistent estimator)。总而言之，由于Y与u相关，估计量b_2是有偏的(对小样本)和不一致的(对大样本)。这正是对联立方程模型使用最小二乘法失效的原因。显然，我们需要探究新的估计方法。我们将在下一节中讨论。如果回归方程中解释变量与该方程中的随机误差项相关，那么这个解释变量也就变成了一个随机变量。在前面考虑的大部分回归模型中，我们或者假设解释变量取固定值，或者 (如果解释变量是随机变量)假设解释变量与随机误差项不相关。在本例中，却不是这两种情况。

在继续新的内容之前，注意方程 (15-7)的一条有趣性质：收入是外生变量投资 I和随机项u的函数。像这样把内生变量表示为外生变量和随机项的方程称为简化方程 (reduced form equation)。随后我们将会看到这种简化方程的作用。

如果将方程(15-7)带入消费函数(15-1)，我们得到了简化形式的消费函数方程：

$$C_t = \frac{B_1}{1-B_2} + \frac{B_2}{1-B_2} I_t + \frac{1}{1-B_2} \mu_t \tag{15-8}$$

方程(15-8)表明：内生变量C仅仅是外生变量I和随机项u的函数。

15.3 间接最小二乘法

从上面的讨论中可以看出，由于Y与u相关，我们不能用普通最小二乘法来估计消费函数中的参数B_1、B_2。那么，还有其他的方法吗？观察方程 (15-8)，为什么不用普通最小二乘法仅仅作C对I的回归呢？的确可以，因为我们假定 I是外生决定的，与u不相关；这与原始的消费函数(15-1)不同。

但是，怎样通过回归方程 (15-8)来估计原始消费函数 (15-1)中的参数呢？这正是我们所关注的问题。其实很简单，我们将方程 (15-8)重写为：

$$C_t = A_1 + A_2 I_t + v_t \tag{15-9}$$

其中，$A_1 = B_1/(1-B_2)$, $A_2 = B_2/(1-B_2)$, $v_t = u_t/(1-B_2)$。

与u一样，v也是随机误差项。A_1、A_2称为简化系数，因为它们是简化模型中的参数。不难发现，简化系数是原始消费函数结构系数的(非线性)组合。

根据上面给出的A，B系数之间的关系，很容易验证：

$$B_1 = \frac{A_1}{1+A_2} \tag{15-10}$$

$$B_2 = \frac{A_2}{1+A_2} \tag{15-11}$$

因此，一旦估计出A_1、A_2，就很容易“推导出”B_1和B_2。

这种估计消费函数(15-1)参数的方法称为间接最小二乘法(ILS)，因为我们首先用普通最小二乘法估计出简化回归方程(15-9)，然后间接地求出原始参数的估计值。间接最小二乘估计量有什么统计性质呢？间接最小二乘估计量是一致估计量，也就是说，随着样本容量的无限增大，间接最小二乘估计量收敛于真实总体值。但是，对于小样本或有限个样本，间接最小二乘估计量可能是有偏的。与间接最小二乘估计量相比，普通最小二乘估计量是有偏的和非一致的。[1]

15.4 实例

我们通过一个具体实例说明间接最小二乘法的运用。表15-1提供了美国1960～1994年消费、收入和投资的数据。以1992年美元价计算(也即，以1992年美元购买力为基准)，单位为亿美元。需要指出的是，为了与凯恩斯的收入决定模型保持一致，收入的数据可简单地看作是消费和投资支出之和。

按照上面讨论的间接最小二乘法，我们首先估计出简化的回归方程(15-8)。利用表15-1提供的数据，得到下面的回归结果；我们用标准形式给出回归结果[形如式(6-48)]：

$$\hat{C}_t = 191.810\ 8 + 4.464\ 07 I_t \qquad (15\text{-}12)$$

$$\text{se}=(130.033\ 0)\quad (0.207\ 1)\qquad r^2=0.933\ 6$$

$$t=(1.475\ 0)\quad (21.553\ 1)$$

表15-1 美国的收入、消费和投资

年份	收　入	个人消费	国内私人总投资
1960	170 3.100	143 2.600	270.500 0
1961	172 6.700	146 1.500	265.200 0
1962	183 2.300	153 3.800	298.500 0
1963	191 4.700	159 6.600	318.100 0
1964	203 6.900	169 2.300	344.600 0
1965	219 1.600	179 9.100	392.500 0
1966	232 5.500	190 2.000	423.500 0
1967	236 5.500	195 8.600	406.900 0
1968	250 0.000	207 0.200	429.800 0
1969	260 1.900	214 7.500	454.400 0
1970	261 7.300	219 7.800	419.500 0
1971	274 6.900	227 9.500	467.400 0
1972	293 8.000	241 5.900	522.100 0
1973	311 6.100	253 2.600	583.500 0
1974	305 9.100	251 4.700	544.400 0
1975	301 0.500	257 0.000	440.500 0
1976	325 0.900	271 4.300	536.600 0
1977	345 6.900	282 9.800	627.100 0
1978	363 7.600	295 1.600	686.000 0
1979	372 4.700	302 0.200	704.500 0
1980	363 5.900	300 9.700	626.200 0
1981	373 6.100	304 6.400	689.700 0
1982	367 1.900	308 1.500	590.400 0
1983	388 8.400	324 0.600	647.800 0

1 详细证明可查阅Damodar N.Gujarati, *Basic Econometrics*, 3d ed., McGraw-Hill, New York, 1995, Chap.18.

(续)

年份	收 入	个人消费	国内私人总投资
1984	423 9.200	340 7.600	831.600 0
1985	439 5.700	356 6.500	829.200 0
1986	452 2.500	370 8.700	813.800 0
1987	464 2.800	382 2.300	820.500 0
1988	479 8.700	397 2.700	826.000 0
1989	492 6.500	406 4.600	861.900 0
1990	494 9.500	413 2.200	817.300 0
1991	484 3.100	410 5.800	737.300 0
1992	501 0.200	421 9.800	790.400 0
1993	519 7.000	433 9.700	857.300 0
1994	545 0.700	447 1.100	979.600 0

资料来源：《总统经济报告》，1996，表B-2，第282页。

注：收入=消费支出+国内私人总投资。所有数据均以1992年美元不变价计算，单位是亿美元。

因而，简化方程(15-8)中的系数A_1，A_2的估计值分别为a_1=191.810 8，a_2=4.464 0。现在，我们根据式(15-10)、(15-11)求得消费函数(15-1)中的参数B_1，B_2的估计值为：

$$b_1 = \frac{a_1}{1+a_2} = \frac{191.8108}{1+4.4640} = 35.104\,4 \tag{15-13}$$

$$b_2 = \frac{a_2}{1+a_2} = \frac{4.4640}{1+4.464\,0} = 0.816\ 9 \tag{15-14}$$

这就是消费函数中参数的间接最小二乘估计值。估计的消费函数为：

$$\hat{C}_t = 35.104\,4 + 0.816\,9Y_t \tag{15-15}$$

因此，估计的MPC约为0.82。

为了比较，我们给出用普通最小二乘法估计得到的回归结果，即直接作C对Y的回归：

$$\hat{C}_t = 6.541\,5 + 0.825\,2Y_t \tag{15-16}$$

$$\text{se}=(23.980\,2)\quad(0.006\,6) \qquad r^2=0.997$$

$$t = (0.272\,7) \qquad (124.603\,1)$$

观察消费函数中的间接最小二乘估计值与普通最小二乘估计值的差别。虽然边际消费倾向没有明显的差别，但是截距却相差很大。究竟哪个结果为真呢？我们知道，在此例中，由于联立问题的存在，用普通最小二乘法得到的结果不仅是有偏的，而且还是不一致的。[1]

看来我们总能够用间接最小二乘法估计联立方程模型中的参数。问题是能否从简化形式的估计值得到原始的结构参数：有些时候可以，有些时候却不能。答案在于模型是否可识别。下一节我们将讨论这一问题，在随后的章节中我们还将介绍估计联立方程模型参数的其他方法。

15.5 模型识别问题

我们回到例15.2的需求和供给模型一例中。假定我们仅仅根据价格P和需求量Q的数据估计需求函数，作Q对P的回归。我们如何知道这个回归确实估计了需求函数呢？你或许会说，

1 注意：我们已经给出了OLS回归方程(15-16)的标准差和t值，但未给出ILS的标准差和t值。这是因为，后者的系数[根据式(15-13)、(15-14)得到]是a_1和a_2的非线性函数，没有一个简单的方法求非线性函数的标准差。

如果估计的斜率为负，它就是需求函数，因为价格和需求量之间呈反向变动关系。但是，如果斜率为正，又会怎样呢？你是否会说它是供给函数呢？因为价格和需求量之间是正向关系。

你会发现在需求量对价格的简单回归中的一个潜在问题：给定一组 P_t 和 Q_t，表示了供给曲线和需求曲线的交点，因为均衡条件是供给等于需求。为了更清楚地看到这一点，我们可以看图15-2。图15-2(a)给出了P、Q的散点图。每一个点都表示了需求曲线和供给曲线的交点，见图(b)。现在我们考虑单独的一个点，见图(c)。在通过该点的一簇曲线中，我们如何确定哪一条是需求曲线，哪一条是供给曲线？显然，需要知道额外的一些关于需求曲线和供给曲线特性的信息。例如，如果由于收入、偏好等因素导致需求曲线的移动，但供给曲线保持相对稳定，见图15-2(d)，这些散点就描绘出了供给曲线。在此情况下，我们说供给曲线是可识别的。也就是说，我们能够惟一地估计出供给曲线的参数。同样地，如果由于天气或其他外生因素导致供给曲线发生移动，但是需求曲线保持相对稳定，见图15-2(e)，这些散点就描绘了需求曲线。在此情况，我们说需求曲线是可识别的；也就是说，我们能够惟一地估计出需求曲线的参数。

因此，识别问题就成为我们能否惟一地估计出某一方程(或是需求或供给函数)的参数。如果能够惟一地估计出参数，那么就称该方程是恰度识别的(exactly identified)。如果我们不能估计出参数，我们就称该模型是不可识别的(unidentified, or underindentified)。有时，方程中的一个或几个参数有若干个估计值，我们就称该方程是过度识别的(overidentified)。下面我们详细讨论每种情况。

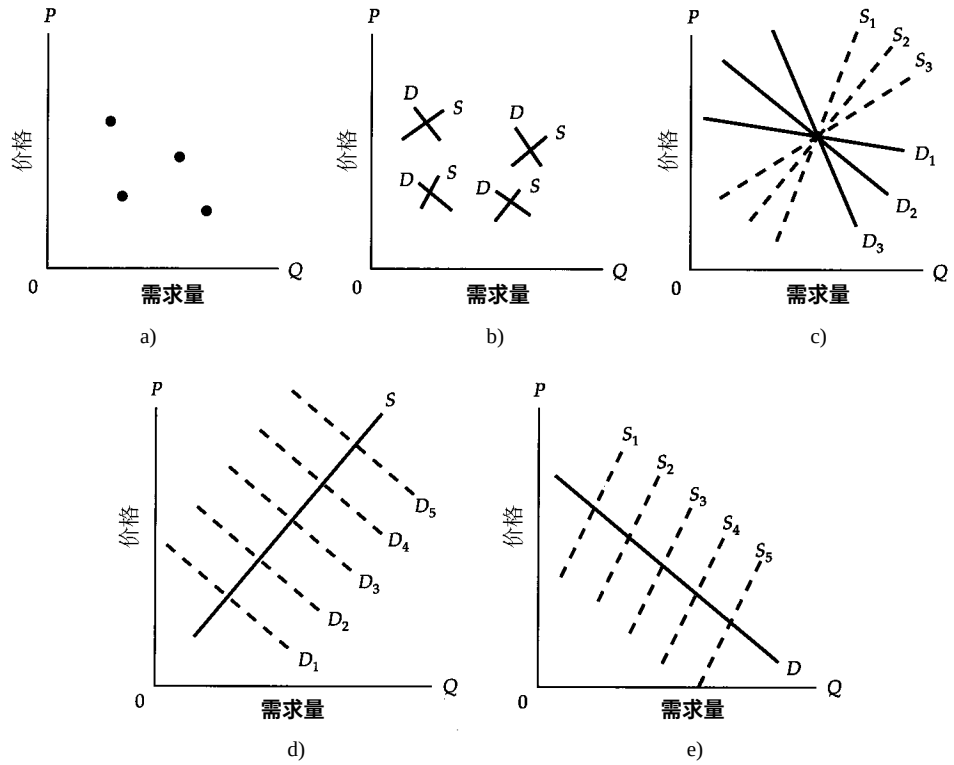


图15-2 供需函数与识别问题

15.5.1 不可识别

我们再来看例15-2。根据均衡条件，供给等于需求，我们得到：

$$A_1+A_2P_t+u_{1t}=B_1+B_2P_t+u_{2t} \tag{15-17}$$

求解方程(15-17)，我们得到均衡价格

$$P_t = \prod_1 + v_{1t} \tag{15-18}$$

其中，

$$\prod_1 = \frac{B_1 - A_1}{A_2 - B_2} \tag{15-19}$$

$$v_{1t} = \frac{\mu_{2t} - \mu_{1t}}{A_2 - B_2} \tag{15-20}$$

其中，v_1是随机误差项，它是u_s的线性组合。符号$\prod$读作“pai”，这里是用来表示简化形式的回归系数。

将式(15-18)中的P_t代到例15.2中的需求或供给方程，我们得到均衡需求量：

$$Q_t = \prod_2 + \mu_{2t} \tag{15-21}$$

其中，

$$\prod_2 = \frac{A_2 B_1 - A_1 B_2}{A_2 - B_2} \tag{15-22}$$

$$v_{2t} = \frac{A_2 \mu_{2t} - B_2 \mu_{1t}}{A_2 - B_2} \tag{15-23}$$

其中，v_2也是随机误差项。

方程(15-19)和(15-21)是简化形式的回归方程。现在需求和供给模型有四个结构参数A_1, A_2, B_1和B_2，但是没有惟一的方法根据$\prod_1$、$\prod_2$来估计这两个简化方程中的参数。根据代数知识我们知道，估计四个未知数必须有四个(独立的)方程。与此同时，如果求解简化方程(15-19)和(15-21)，我们会发现没有解释变量，只有常量$\prod'$，这些常量仅仅给出了价格和需求量的平均值(为什么？)，通过这两个平均值我们根本无法求出四个结构参数。简言之，需求和供给函数都是不可识别的。

15.5.2 恰度识别

前面讨论了用间接最小二乘法估计凯恩斯消费函数的情况，根据简化形式的回归方程(15-12)，我们可以获得消费函数参数的惟一值[参见式(15-3)、(15-4)]。

为了进一步说明恰度识别，我们仍以需求和供给模型为例，但是现在我们对模型做如下修正：

需求函数：$Q_t^d = A_1 + A_2 P_t + A_3 X_t + \mu_{1t}$ (15-24)

供给函数：$Q_t^d = B_1 + B_2 P_t + \mu_{2t}$ (15-25)

其中，增加的变量X定义为消费者的收入。因而，需求函数表明，需求量是价格和消费者收入的函数；需求理论一般将价格和收入作为决定需求量的两个主要决定因素。把收入变量包括进模型中将会对消费者行为提供额外的信息。这里假定了消费者收入是内生决定的。

根据市场出清机制，需求量=供给量，得到：

$$A_1 + A_2 P_t + A_3 X_t + u_{1t} = B_1 + B_2 P_t + u_{2t} \tag{15-26}$$

求解式(15-26)得到均衡的P_t值如下：

$$P_t = \prod_1 + \prod_2 X_t + v_{1t} \tag{15-27}$$

简化形式的系数为：

$$\prod_1 = \frac{B_1 - A_1}{A_2 - B_2} \tag{15-28}$$

$$\prod_2 = \frac{A_3}{A_2 - B_2} \tag{15-29}$$

$$v_{it} = \frac{\mu_{2t} - \mu_{1t}}{A_2 - B_2} \tag{15-30}$$

将均衡的P_t值代入上述的需求和供给函数，我们得到市场出清或均衡的需求量：

$$Q_t = \prod_3 + \prod_4 X_t + v_{2t} \tag{15-31}$$

其中，

$$\prod_3 = \frac{A_3B_1 - A_1B_2}{A_2 - B_2} \tag{15-32}$$

$$\prod_4 = \frac{A_3B_2}{A_2 - B_2} \tag{15-33}$$

$$v_{2t} = \frac{A_2\mu_{2t} - B_2\mu_{1t}}{A_2 - B_2} \tag{15-34}$$

由于式(15-27)、(15-31)都是简化形式的回归方程，因此，用普通最小二乘法可以估计它们的参数。问题是我们能否从简化形式系数中惟一地估计出结构方程中的参数呢？

观察需求和供给模型(15-24)、(15-25)，共包括五个结构参数，A_1,A_2,A_3,B_1和B_2。但是我们仅有四个方程，即四个简化形式的系数(四个$\prod$)。因此，我们无法求得所有五个结构系数的惟一值。但哪一个系数值可以惟一确定呢？读者可以验证，供给函数的参数能够惟一确定：

$$B_1 = \prod_3 - B_2 \prod_1 \tag{15-35}$$

$$B_2 = \frac{\prod_4}{\prod_2} \tag{15-36}$$

因此，供给函数是恰度识别的。但需求函数却是不可识别的，因为我们无法估计其参数——系数A。

一个有趣的事实是：正是由于需求函数中这个增加的变量才使得我们能够识别供给函数。为什么呢？把收入变量包括进需求函数为我们提供了一些关于函数变动的额外信息，参见图15-2(d)。该图表明，稳定的供给曲线与移动的需求曲线的交点是如何使我们能够识别供给曲线的。

怎样使需求函数可识别呢？假定我们把滞后一期的价格P_{t-1}作为增加变量进入到供给函数(15-25)之中。也就是说供给不仅依赖于当期价格，而且还依赖于上期价格。对许多农产品而言，这完全是不合理的假定。由于在t期，已经知道P_{t-1}的值，我们可把它看作是外生变量或预定变量。因而，新的模型为：

需求函数：$Q_t^d = A_1 + A_2P_t + A_3X_t + u_{1t}$ (15-37)

供给函数：$Q_t^d = B_1 + B_2P_t + B_3P_{t-1} + u_{2t}$ (15-38)

在市场出清的条件下，利用式(15-37)、(15-38)，读者可以求得简化形式的回归方程并能验证此时的需求和供给函数都是可识别的；每一个简化方程包括X_t和P_t两个解释变量，由于这两个变量都是由外生决定的，因而它们与随机误差项不相关。再一次提醒注意：变量的进入或排除是如何帮助我们识别模型的，也即如何得到惟一的参数值。因而，将变量P_{t-1}排除在需求函数之外使得我们能够识别需求函数，正如将变量X_t(收入)排除在供给函数之外一样，也是为了使我们能够识别供给函数。可以得到一个推论：联立方程系统中，如果一个方程包含了系统内的所有变量(内生的和外生的)，那么它将不能被识别。随后我们将给出一个简单的识别规则来概括这一思想(参见15.6节)。

15.5.3 过度识别

虽然从方程中排除某些变量或许能够使我们可以识别它，但是有时会出现过度的情况，从而导致了过度识别问题。即模型中的某一方程的参数可能有不止一个的估计值。我们来看一下这种情况是如何发生的。

我们仍以需求-供给模型为例：

需求函数：$Q_t=A_1+A_2P_t+A_3X_t+A_4W_t+u_{1t}$ (15-39)

供给函数：$Q_t=B_1+B_2P_t+B_3P_{t-1}+u_{2t}$ (15-40)

其中，增加的变量W_t表示消费者财富。对于许多商品，收入和财富是决定需求的重要因素。将(15-37)、(15-38)的需求和供给模型与(15-39)和(15-40)相比较。在(15-37)供给函数仅仅排除了收入变量，而(15-39)供给函数不仅排除了收入变量还排除了财富变量。从供给函数中排除收入变量使得我们能够识别它，现在从供给函数中将收入和财富都排除出去，将导致过度识别，即供给参数B_2有两个估计值，证明如下：

令式(15-39)与(15-40)相等，可得下面的简化方程：

$$P_t=\prod_1+\prod_2X_T+\prod_3W_t+\prod_4P_{t-1}+v_{1t} \tag{15-41}$$

$$Q_t=\prod_5+\prod_6X_T+\prod_7W_t+\prod_8P_{t-1}+v_{2t} \tag{15-42}$$

其中，

$$\prod_1=\frac{B_1-A_1}{A_2-B_2} \qquad \prod_2=-\frac{A_3}{A_2-B_2}$$

$$\prod_3=\frac{A_4}{A_2-B_2} \qquad \prod_4=\frac{B_3}{A_2-B_2}$$

$$\prod_5=\frac{A_2B_1-A_1B_2}{A_2-B_2} \qquad \prod_6=\frac{A_3B_2}{A_2-B_2} \tag{15-43}$$

$$\prod_7=\frac{A_4B_2}{A_2-B_2} \qquad \prod_8=\frac{A_2B_3}{A_2-B_2}$$

$$v_{1t}=\frac{\mu_{2t}-\mu_{1t}}{A_2-B_2} \qquad v_{2t}=\frac{A_2\mu_{2t}-B_2\mu_{1t}}{A_2-B_2}$$

在需求供给模型中，我们共考虑了七个结构系数，其中四个 A，三个B。但在简化模型(15-43)中共有八个系数。即方程的个数比求知参数多，所以很显然参数有不止一个的解。读者很容易验证，事实上，B_2有两个值：

$$B_2=\frac{\prod_7}{\prod_3} \quad \text{或} \quad B_2=\frac{\prod_6}{\prod_2} \tag{15-44}$$

我们没有任何理由认为这两个估计值是相等的。

由于式(15-43)中的所有简化方程系数的分母中都包含了 B_2，因此B_2的不确定将导致其他结构参数的不确定。为什么会产生这个结果呢？看来是由于有太多的信息 —— 排除收入变量或财富变量的其中一个就足以识别供给方程了。这是不可识别的相反情况，在那里，信息又太少了。这里的一个要点是：并非信息越多越好！需特别指出的是：过度识别问题并不是因为随意增加变量而发生的。有些时候经济理论会告诉我们哪些变量应从模型之中排除，哪些变量应包括到模型之中，从而判定模型是不可识别的还是可识别的 (恰度识别或过度识别)。

总之，联立方程模型中的方程或是不可识别的，或是恰度识别的，或是过度识别的！对于

不可识别的模型我们无能为力，只有假定它是正确的。不可识别不是一个可以通过增大样本容量而得到解决的统计问题。你能观察图 15-2(a)中每一年的四个点的图形，但是你无法知道由这四个点所生成的需求和供给曲线的斜率。如果方程是恰度识别的，我们可利用间接最小二乘法来估计其参数。如果方程是过度识别的，用间接最小二乘法将得不到惟一的参数值。幸运的是，我们可以利用两阶段最小二乘法 (2SLS)来估计过度识别方程的参数。在介绍两阶段二乘法之前，我们先来看是否存在系统的方法判定方程是不可识别的，恰度识别的或是过度识别的。用简化方程来判定识别问题相当麻烦，尤其是当模型包括若干个方程的时候。

15.6 模型识别的判定规则：识别的阶条件

为了理解阶条件，我们先介绍下面的符号：

m —— 模型中内生变量(联合相关)的个数

k —— 不包括在该方程中的所有变量(内生变量和外生变量)的个数。

则，

(1) 若$k=m-1$，方程恰度识别。

(2) 若$k>m-1$，方程过度识别。

(3) 若$k<m-1$,方程不可识别。

运用阶条件，我们只需数清楚模型中内生变量的个数 (=模型中方程的个数)以及不包括在方程中的变量的总数 (内生的和外生的)。虽然，识别的阶条件仅仅是必要条件而非充分条件，但在实际中还是非常有用的。

对式(15-39)和(15-40)的供给和需求模型用阶条件，此时 $m=2$，供给函数不包括变量 X_t和 W_t(即$k=2$)。由于$k>m-1$，所以供给方程是过度识别的。需求函数不包括 P_{t-1}，由于$k=m-1$，所以需求方程是恰度识别的。但这只是较为简单的情况。如果我们想根据简化模型的系数 (15-43)来估计需求函数的参数，则估计值不是惟一的，因为进入计算的 B_2有两个取值[见式(15-43)]。但用两阶段最小二乘法就可避免出现这种情况。

15.7 对过度识别方程的估计：两阶段最小二乘法

我们考虑下面的模型：

收入函数：$Y_t=A_1+A_2M_t+A_3I_t+A_4G_t+u_{1t}$ (15-45)

货币供给函数：$M_t=B_1+B_2Y_t+u_{2t}$ (15-46)

其中 Y —— 收入；

M —— 货币；

I —— 投资支出；

G —— 政府的商品和劳物支出；

u_1, u_2 —— 随机误差项。

在这个模型中，I和G是外生给定的。

根据数量论和凯恩斯收入决定理论，收入函数表明了收入是由货币供给、投资支出和政府支出决定的。货币供给函数表明，货币供给是联邦储备银行根据收入水平决定的。显然，这是一个联立问题，因为收入和货币供给之间存在着反馈。

利用阶条件，我们可以判定收入方程是不可识别的 (因为它包括了所有变量)，但货币供给方程却是过度识别的，因为它排除了系统内的两个变量 (注：在这个模型中，$m=2$)。

由于收入方程是不可识别的，因此我们无法估计其参数。货币供给方程又怎样呢？由于它是过度识别的，所以若用间接最小二乘法估计其参数，则不能得到惟一的估计值；事实上，B_2确有两个值。如果用普通最小二乘法又如何呢？由于Y与随机误差项u_2之间可能相关，根据我们前面的讨论，OLS估计值将是不一致的。那么，我们将作何选择呢？

假定在货币供给函数中，我们找到一个替代变量(surrogate or proxy)或工具变量(instrumental)来代替Y，这个变量虽与Y相像，但却与u_2不相关。若能得到这样一个变量，那么就可直接用普通最小二乘法来估计货币供给函数中的参数了。(为什么？)但是如何得到这样的一个工具变量呢？答案是用两阶段最小二乘法。正如其名，这种方法分作两个阶段应用普通最小二乘法：

第一阶段，首先作Y对每个预定变量的回归(注意是整个模型中的所有预定变量，而并非只是该方程的)以剔除Y与随机误差项u_2之间可能存在的相关因素。在此例中，就是作Y对I(国内私人总投资)和G(政府支出)的回归：

$$Y_t = \prod_1 + \prod_2 I_t + \prod_3 G_t + w_t \tag{15-47}$$

其中，w是随机误差项。根据式(15-47)，得到：

$$\hat{Y}_t = \prod_1 + \prod_2 I_t + \prod_3 G_t \tag{15-48}$$

其中，$\hat{Y}_t$是在给定I和G值条件下的Y的估计值。注意，$\prod$系数上的∧表明了它们是真实的$\prod$的估计值。

因此，我们可以将式(15-47)写为：

$$Y_t = \hat{Y}_t + w_t \tag{15-49}$$

式(15-49)表明(随机的)Y由两部分组成：$\hat{Y}_t$[根据式(15-47)，它是预定变量I和G的线性组合]和随机部分w_t。根据最小二乘理论，$\hat{Y}_t$与w_t不相关(为什么？参见习题5.13)。

第二阶段，过度识别的货币供给函数现在可写为：

$$\begin{aligned} M_t &= B_1 + B_2(\hat{Y}_t + w_t) + u_{2t} \\ &= B_1 + B_2\hat{Y}_t + (u_{2t} + B_2 w_t) = B_1 + B_2\hat{Y}_t + v_t \end{aligned} \tag{15-50}$$

其中，$v_t = u_{2t} + B_2 w_t$

比较方程(15-50)和(15-46)，形式很相像，所不同的是用$\hat{Y}_t$代替Y_t，这种形式有什么优点呢？能够证明，虽然在原始的货币供给函数(15-46)中Y可能与随机误差项u_2相关，(因而导致OLS失效)，但是式(15-50)中的$\hat{Y}$与v_t却是渐进无关的(对于大样本，或更准确地，随着样本容量无限增大)。因而可对式(15-50)用普通最小二乘法。这是对(15-46)式直接用普通最小二乘法的改进，因为在那种情况下，估计值可能是有偏的和不一致的。[1]

15.8 2SLS：一个数字例子

继续货币供给和收入模型一例。表15-2给出了Y(收入，用GDP来度量)、M(货币供给，用M_2度量)、I(投资，用国内私人总投资，GPDI，来度量)和G(联邦政府支出)的数据。单位为亿美元。利率以百分比形式给出(用6月期国债利率来度量)。所有的数据是年度数据，即从1970年到1994年。

1 进一步的讨论参见Damodar N.Gujarati, Basic Econometrics, 3d ed., McGraw-Hill, New York, 1995.Chap.20.

表15-2 美国宏观经济数据

时间	GDP	GDPI	Government	M2	TB
1970	103 5.600	150.200 0	115.900 0	628.100 0	6.562 000
1971	112 5.400	176.000 0	117.100 0	712.700 0	4.511 000
1972	123 7.300	205.600 0	125.100 0	805.200 0	4.466 000
1973	138 2.600	242.900 0	128.200 0	861.000 0	7.178 000
1974	149 6.900	245.600 0	139.900 0	908.500 0	7.926 000
1975	163 0.600	225.400 0	154.500 0	102 3.700	6.122 000
1976	181 9.000	286.600 0	162.700 0	116 3.700	5.266 000
1977	202 6.900	356.600 0	178.400 0	128 6.500	5.510 000
1978	229 1.400	430.800 0	194.400 0	138 8.600	7.572 000
1979	255 7.500	480.900 0	215.000 0	149 6.900	10.01 700
1980	278 4.200	465.900 0	248.400 0	162 9.300	11.37 400
1981	311 5.900	556.200 0	284.100 0	179 3.300	13.77 600
1982	324 2.100	501.100 0	313.200 0	195 3.200	11.08 400
1983	351 4.500	547.100 0	344.500 0	218 7.700	8.750 000
1984	390 2.400	715.600 0	372.600 0	237 8.400	9.800 000
1985	418 0.700	715.100 0	410.100 0	257 6.000	7.660 000
1986	442 2.200	722.500 0	435.200 0	282 0.300	6.030 000
1987	469 2.300	747.200 0	455.700 0	292 2.300	6.050 000
1988	504 9.600	773.900 0	457.300 0	308 3.500	6.920 000
1989	543 8.700	829.200 0	477.200 0	324 3.000	8.040 000
1990	574 3.800	799.700 0	503.600 0	335 6.000	7.470 000
1991	591 6.700	736.200 0	522.600 0	345 7.900	5.490 000
1992	624 4.400	790.400 0	528.000 0	351 5.300	3.570 000
1993	655 0.200	871.100 0	522.100 0	358 3.600	3.140 000
1994	693 1.400	101 4.400	516.300 0	361 7.000	4.660 000

注：GDP：国内总产出，亿美元，表B-1，第280页
GDPI：国内私人总投资，亿美元，表B-1，第280页
Government：联邦政府支出，亿美元，表B-1，第281页
M2：货币供给，表B-65，第355页
TB：6月期国债利率，表B-65，第360页
资料来源：《总统经济报告》，1996

阶段1回归。为了估计货币供给函数(15-46)的参数，我们首先作Y(收入)对变量I和G的回归，这里I和G是外生决定的。回归结果如下：

$$\hat{Y}_t = -333.178\ 9 + 2.030\ 7I_t + 8.718\ 8G_t$$
$$\text{se} = (151.174\ 1)\quad (1.005\ 4)\quad (1.655\ 0)\qquad R^2 = 0.974\ 2 \qquad (15\text{-}51)$$
$$t = (-2.203\ 9)\quad (2.019\ 8)\quad (5.268\ 2)$$

读者可对回归结果作出解释。注意，所有的系数在5%的显著水平上都是统计显著的。

阶段2回归。估计货币供给函数，作M对估计的Y(15-48)的回归，而不是对原始Y的回归。结果如下：

$$\hat{M}_t = 115.032\ 7 + 0.560\ 5\hat{Y}_t$$
$$\text{se} = (54.402\ 1)\quad (0.013\ 6)\qquad r^2 = 0.986\ 3 \qquad (15.52)^1$$
$$t = (2.114\ 4)\quad (40.973\ 5)$$

1 这些标准差正确地反映了随机误差项 v_t的性质。感兴趣的同学可参阅 Damodar N.Gujarati, *Basic Econometrics*, 3d ed., McGraw-Hill, New York, 1995,p. 705。

注：右边的Y上有一“ ^。”

OLS回归。为了比较，我们给出OLS的(显然，这里用OLS法是不适合的)回归结果：

$$\hat{M}_t = 146.817\ 9 + 0.551\ 5\ \hat{Y}_t$$
$$\text{se} = (53.323\ 5) \quad (0.013\ 3) \quad r^2 = 0.986\ 6 \tag{15-53}$$
$$t = (2.753\ 3) \quad (41.245\ 8)$$

比较2SLS和OLS的回归结果，你可能会说回归结果没有很大的差别。在这个例子中可能如此，但是不能保证所有的情况都如此。除此之外，我们知道，从理论上说， 2SLS比OLS要好，尤其是对大样本。

在结束讨论联立方程模型之前，还需提及的是：除了 2SLS和OLS之外还有其他的估计模型的方法(完全信息的最大概似法， the method of full information maximum likehood)。但对这些方法的讨论已超出本书的范围。[1] 本章我们主要的目的是介绍联立方程模型的基本框架，让读者知道我们遇到的不仅仅是单方程回归模型。

15.9 小结

与前面几章讨论的单方程模型相比，对于联立方程模型，变量在一个方程中是(内生的)应变量，而在另一个方程中却以解释变量的形式出现。因而，变量之间存在着反馈。这种反馈就产生了联立问题，使得用OLS法在估计单方程参数时失效。这是因为以解释变量形式出现的内生变量在另一个方程中可能与该方程中的随机误差项相关。这就违背了使用OLS法的假定之一：解释变量要么是固定的或非随机的，要么是随机的，但与误差项不相关。正是由于这个原因，使用OLS法得到的估计值将是有偏的和不一致的。

除了联立问题之外，联立方程模型还有识别问题。识别问题是指我们不能惟一地估计方程中的参数。因此，在估计联立方程模型之前，必须判定模型中的方程是否可识别。

一种判定模型是否可识别的较烦琐的方法是利用模型的简化形式。简化方程描述了应变量仅仅是外生变量或预定变量 (即变量的值由模型之外决定)的函数。如果在简化形式的参数与原始方程的参数之间一一对应，那么原始的方程就是可识别的。

判定模型是否可识别的一种简单方法是利用阶条件。阶条件就是数清模型中方程的个数以及模型中变量的个数(内生变量和外生变量之和)。然后根据方程中排除的变量 (但包括在模型其他方程中)的个数，利用阶条件判定模型是不可识别的、恰度识别还是过度识别的。如果不能估计出参数值，则该方程是不可识别的；如果能够得到参数的惟一值，则该方程是恰度识别的；另一方面，如果方程中的一个或多个参数有不止一个的估计值，则该方程是过度识别的。

如果方程是不可识别的，那么我们也就无能为力了。我们只有改变模型的假定 (也就是说，建立一个新的模型)。如果模型是恰度识别的，利用间接最小二乘法 (ILS)能够估计该方程，然后根据简化形式的参数求出原始方程的参数。间接最小二乘估计量是一致的，即随着样本容量的无限增加，估计量将收敛于其真实值。过度识别方程的参数可利用两阶段最小二乘法 (2SLS)估计。其基本思想是用一个与随机误差项不相关的变量代替与误差项相关的解释变量。这样的一个变量称为替代变量或工具变量。与 ILS估计量一样，2SLS估计量也是一致估计量。

习题

15.1 什么是联立问题？

1 有兴趣的读者可参阅William H.Greene, *Econometric Analysis*, 3d ed., Prentice-Hall, New Jersey, 1997, Chap. 16。

15.2 什么是内生变量和外生变量？

15.3 为什么OLS方法一般不适用于估计联立方程模型中的单方程？

15.4 如果用OLS方法估计联立方程模型中的方程，结果会怎样？

15.5 简化形式的方程有什么作用？

15.6 什么是结构或行为方程？

15.7 什么是间接最小二乘法？什么时候使用它？

15.8 什么是识别问题？为什么它是重要的？

15.9 识别的阶条件是什么？

15.10 为什么说识别的阶条件是必要而非充分条件？

15.11 解释概念：(1)不可识别 (2)恰度识别 (3)过度识别

15.12 如何估计不可识别方程？

15.13 用什么方法估计恰度识别的方程？

15.14 2SLS用于估计那类方程？

15.15 2SLS也能用于估计恰度识别的方程吗？

15.16 考虑下面的双方程模型：

$$Y_{1t}=A_1+A_2Y_{2t}+A_3X_{1t}+u_{1t}$$
$$Y_{2t}=B_1+B_2Y_{1t}+B_3X_{2t}+u_{2t}$$

其中，Y是内生变量，X是外生变量，u是随机误差项。

(a) 求简化形式回归方程？

(b) 判定哪个方程是可识别的。

(c) 对可识别方程，你将用那种方法进行估计，为什么？

(d) 假定，先验地知道$A_3=0$。那么你将如何回答上述问题，为什么？

15.17 考虑下面的模型：

$$Y_{1t}=A_1+A_2Y_{2t}+A_3X_{1t}+u_{1t}$$
$$Y_{2t}=B_1+B_2Y_{1t}+u_{2t}$$

其中，Y是内生变量，X是外生变量，u是随机误差项。根据这个模型，得到简化形式的回归模型如下：

$$Y_{1t}=6+8X_{1t}$$
$$Y_{2t}=4+12X_{1t}$$

(a) 从这些简化方程中，你能估计出那些结构参数？

(b) 如果，先验地知道(1) $A_2=0$ (2) $A_1=0$，你的答案将作何修改？

15.18 考虑下面的模型：

$$R_t=A_1+A_2M_t+A_3Y_t+u_{1t}$$
$$Y_t=B_1+B_2R_t+u_{2t}$$

其中，Y=收入(用GDP来度量)，R=利率(用6月期国债利率，%)，M=货币供给(用$M2$度量)。假定M外生给定。

(a) 该模型之后的经济原理是什么？(提示：可参阅有关宏观经济学的教科书)

(b) 上述方程可识别吗？

(c) 利用表15-2给出的数据，估计可识别方程的参数。

15.19 考虑下面改进过的习题15.18的模型：

$$R_t=A_1+A_2M_t+A_3Y_t+u_{1t}$$
$$Y_t=B_1+B_2R_t+B_3I_t+u_{2t}$$

I为增加的变量，代表投资(用国内私人总投资来度量)。假定M和I外生给定。

(a) 上述方程哪一个是可识别的？

(b) 利用表15-2提供的数据，估计可识别方程的参数。

(c) 你对本题和上一题的不同结果有什么看法？

附录15A OLS估计量的不一致性

为了证明OLS估计量b_2是B_2的不一致估计量（由于Y_t与u_t相关），我们先看OLS估计量(15-6)：

$$b_2 = \frac{\sum (C_t - \bar{C})(Y_t - \bar{Y})}{\sum (Y_t - \bar{Y})^2} = \frac{\sum C_t y_t}{\sum y_t^2} \tag{15A-1}$$

其中，$y_t = (Y_t - \bar{Y})$

现将式(15-1)带入，得：

$$b_2 = \frac{\sum (B_1 + B_2 Y_t + \mu_t) y_t}{\sum y_t^2} = B_2 + \frac{\sum y_t \mu_t}{\sum y_t^2} \tag{15A-2}$$

其中最后一步用到，$\sum y_t = 0$，$\sum Y_t y_t / \sum y_t^2 = 1$（为什么？）

对式(15A-2)取期望：

$$E(b_2) = B_2 + E\left[\frac{\sum y_t \mu_t}{\sum y_t^2}\right] \tag{15A-3}$$

不幸的是，我们不能轻易地估计出第二项的期望，因为期望的运算符E是一个线性运算符。[注：$E(A/B) \neq E(A)/E(B)$]直观上很清楚，除非式(15A-3)中的第二项为零，否则b_2是B_2的有偏估计量。

b_2不仅是有偏的，而且还是不一致的。估计量称为一致估计量，如果它的概率极限等于其真实(总体)值。[1] 运用概率极限的性质[2]，可以得到：

$$\text{plim}(b_2) = \text{plim}(B_2) + \text{plim}\left[\frac{\sum y_t \mu_t}{\sum y_t^2}\right] = B_2 + \text{plim}\left[\frac{\sum y_t \mu_t / n}{\sum y_t^2 / n}\right]$$

$$= B_2 + \frac{\text{plim}(\sum y_t \mu_t / n)}{\text{plim}(\sum y_t^2 / n)} \tag{15A-4}$$

这里用了概率极限的性质——常数的概率极限等于常数本身,比值的概率极限等于极限概率的比值。

随着n的无限增大，可以证明：

$$\text{plim}(b_2) = B_2 + \frac{1}{1 - B_2}\left(\frac{\sigma^2}{\sigma_y^2}\right) \tag{15A-5}$$

其中，σ^2是u的方差，σ_y^2是Y的方差。

因为B_2(MPC)位于0、1之间，且方程(15A-5)中的两个方差项均为正，所以显然b_2的概率极限比B_2大，也就是说，b_2过高的估计了B_2，而且无论样本容量如何，这个偏差总是存在的。

---

1 如果$\lim_{n\to\infty} P\{|b_2 - B_2| < d\} = 1$，其中$d > 0$，$n$是样本容量，我们就称$b_2$是$B_2$的估计量，简写为，$n \to \infty$ $\text{plim}(b_2) = B_2$。更详细的内容参见 Damodar N.Gujarati, *Basic Econometrics*, 3d ed.,McGraw-Hill, New York, 1995, pp.782-783。

2 虽然$E(A/B) \neq E(A)/E(B)$，但是，$\text{plim}(A/B) = \text{plim}(A)/\text{plim}(B)$.

附 录 A

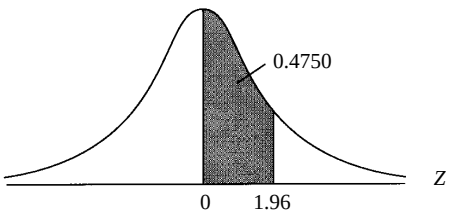
统 计 表

表A-1　标准正态分布下的面积
表A-2　*t*分布的百分数点
表A-3　*F*分布的上百分点
表A-4　X^2分布的上百分点
表A-5　杜宾-瓦尔森*d*统计量：在显著水平0.05与0.01上的d_L与d_U的显著点
表A-6　游程检验中的各游程的临界值

例

Pr (0≤Z≤1.96)=0.4750

Pr (Z≥1.96)=0.5-0.4750=0.025



表A-1　标准正态分布下的面积

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.000 0	0.004 0	0.008 0	0.012 0	0.016 0	0.019 9	0.023 9	0.027 9	0.031 9	0.035 9
0.1	0.039 8	0.043 8	0.047 8	0.051 7	0.055 7	0.059 6	0.063 6	0.067 5	0.071 4	0.075 3
0.2	0.079 3	0.083 2	0.087 1	0.091 0	0.094 8	0.098 7	0.102 6	0.106 4	0.110 3	0.114 1
0.3	0.117 9	0.121 7	0.125 5	0.129 3	0.133 1	0.136 8	0.140 6	0.144 3	0.148 0	0.151 7
0.4	0.155 4	0.159 1	0.162 8	0.166 4	0.170 0	0.173 6	0.177 2	0.180 8	0.184 4	0.187 9
0.5	0.191 5	0.195 0	0.198 5	0.201 9	0.205 4	0.208 8	0.212 3	0.215 7	0.219 0	0.222 4
0.6	0.225 7	0.229 1	0.232 4	0.235 7	0.238 9	0.242 2	0.245 4	0.248 6	0.251 7	0.254 9
0.7	0.258 0	0.261 1	0.264 2	0.267 3	0.270 4	0.273 4	0.276 4	0.279 4	0.282 3	0.285 2
0.8	0.288 1	0.291 0	0.293 9	0.296 7	0.299 5	0.302 3	0.305 1	0.307 8	0.310 6	0.313 3
0.9	0.315 9	0.318 6	0.321 2	0.323 8	0.3264	0.328 9	0.331 5	0.334 0	0.336 5	0.338 9
1.0	0.341 3	0.343 8	0.346 1	0.348 5	0.350 8	0.353 1	0.355 4	0.357 7	0.359 9	0.362 1
1.1	0.364 3	0.366 5	0.368 6	0.370 8	0.372 9	0.374 9	0.377 0	0.379 0	0.381 0	0.383 0
1.2	0.384 9	0.386 9	0.388 8	0.390 7	0.392 5	0.394 4	0.396 2	0.398 0	0.399 7	0.401 5
1.3	0.403 2	0.404 9	0.406 6	0.408 2	0.409 9	0.411 5	0.413 1	0.414 7	0.416 2	0.417 7
1.4	0.419 2	0.420 7	0.422 2	0.423 6	0.4251	0.426 5	0.427 9	0.429 2	0.430 6	0.431 9
1.5	0.433 2	0.434 5	0.435 7	0.437 0	0.438 2	0.439 4	0.440 6	0.441 8	0.442 9	0.444 1
1.6	0.445 2	0.446 3	0.447 4	0.448 4	0.449 5	0.450 5	0.451 5	0.452 5	0.453 5	0.454 5
1.7	0.445 4	0.456 4	0.457 3	0.458 2	0.459 1	0.459 9	0.460 8	0.461 6	0.462 5	0.463 3
1.8	0.464 1	0.464 9	0.465 6	0.466 4	0.467 1	0.467 8	0.468 6	0.469 3	0.469 9	0.470 6
1.9	0.471 3	0.471 9	0.472 6	0.473 2	0.473 8	0.474 4	0.475 0	0.475 6	0.476 1	0.476 7
2.0	0.477 2	0.477 8	0.478 3	0.478 8	0.479 3	0.479 8	0.480 3	0.480 8	0.481 2	0.481 7
2.1	0.482 1	0.482 6	0.483 0	0.483 4	0.483 8	0.484 2	0.484 6	0.485 0	0.485 4	0.485 7
2.2	0.486 1	0.486 4	0.486 8	0.487 1	0.487 5	0.487 8	0.488 1	0.488 4	0.488 7	0.489 0
2.3	0.489 3	0.489 6	0.489 8	0.490 1	0.490 4	0.490 6	0.490 9	0.491 1	0.491 3	0.491 6
2.4	0.491 8	0.492 0	0.492 2	0.492 5	0.492 7	0.492 9	0.493 1	0.493 2	0.493 4	0.493 6

(续)

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.5	0.493 8	0.494 0	0.494 1	0.494 3	0.494 5	0.494 6	0.494 8	0.494 9	0.495 1	0.495 2
2.6	0.495 3	0.495 5	0.495 6	0.495 7	0.495 9	0.496 0	0.496 1	0.496 2	0.496 3	0.496 4
2.7	0.496 5	0.496 6	0.496 7	0.496 8	0.496 9	0.497 0	0.497 1	0.497 2	0.497 3	0.497 4
2.8	0.497 4	0.497 5	0.497 6	0.497 7	0.497 7	0.497 8	0.497 9	0.497 9	0.498 0	0.498 1
2.9	0.498 1	0.498 2	0.498 2	0.498 3	0.498 4	0.498 4	0.498 5	0.498 5	0.498 6	0.498 6
3.0	0.498 7	0.498 7	0.498 7	0.498 8	0.498 8	0.498 9	0.498 9	0.498 9	0.499 0	0.499 0

注：本表给出了正态分布的右侧面积 (也即$Z \geq 0$)。但是由于正态分布关于$Z=0$对称，所以左侧面积与右侧面积相等。例如，$P(-1.96 \leq Z \leq 0)=0.475\ 0$。因此，$P(-1.96 \leq Z \leq 1.96)=2(0.475\ 0)=0.95$。

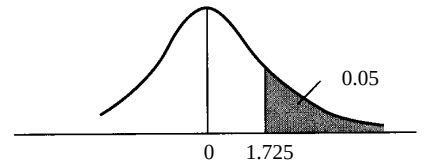
表A-2 *t*分布的百分数点

例:

Pr (t>2.086)=0.025

Pr (t>1.725)=0.05 **自由度**=20

Pr(| t |>1.725)=0.10



Pr	0.25	0.10	0.05	0.025	0.01	0.005	0.001
自由度	0.50	0.20	0.10	0.05	0.02	0.010	0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

注：每栏上端所示较小概率为单侧面积；下端较大的概率为双侧面积。

资料来源： E.S. Pearson**和**H.O.Hartley,eds, *Biometrika Tables for Statisticians*, vol.1,3ded.,table 12,Cambridge University Press, New York, 1996. Reproduced by permission of the editors and trustees of *Biometrika.*

表A-3 *F*分布的

例:

Pr (*F*＞1.59)=0.25

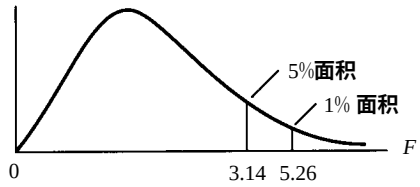
Pr (*F*＞2.42)=0.10 N_1=10

Pr (*F*＞3.14)=0.05 N_2=9

Pr (*F*＞5.26)=0.01

分母N_2的自由度	Pr	分子N_1的自由度 1	2	3	4	5	6	7	8	9	10	11	12
	0.25	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.36	9.41
1	0.10	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	60.5	60.7
	0.05	161	200	216	225	230	234	237	239	241	242	243	244
	0.25	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.39
2	0.10	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.40	9.41
	0.05	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4
	0.01	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4
	0.25	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.45
3	0.10	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.22
	0.05	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
	0.01	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	27.1
	0.25	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08
4	0.10	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.91	3.90
	0.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91
	0.01	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.4
	0.25	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89
5	0.10	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.28	3.27
	0.05	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.71	4.68
	0.01	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.96	9.89
	0.25	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.77
6	0.10	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.92	2.90
	0.05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
	0.01	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
	0.25	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.69	1.69	1.69	1.68
7	0.10	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.68	2.67
	0.05	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57
	0.01	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47
	0.25	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.63	1.62
8	0.10	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.52	2.50
	0.05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
	0.01	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67
	0.25	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58
9	0.10	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.40	2.38
	0.05	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
	0.01	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11

上百分点



分子N_1的自由度													分母N_2的自由度
15	20	24	30	40	50	60	100	120	200	500	∞	Pr	
9.49	9.58	9.63	9.67	9.71	9.74	9.76	9.78	9.80	9.82	9.84	9.85	0.25	
61.2	61.7	62.0	62.3	62.5	62.7	62.8	63.0	63.1	63.2	63.3	63.3	0.10	1
246	248	249	250	251	252	252	253	253	254	254	254	0.05	
3.41	3.43	3.43	3.44	3.45	3.45	3.46	3.47	3.47	3.48	3.48	3.48	0.25	
9.42	9.44	9.45	9.46	9.47	9.47	9.47	9.48	9.48	9.49	9.49	9.49	0.10	2
19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	0.05	
99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	0.01	
2.46	2.46	2.46	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47	0.25	
5.20	5.18	5.18	5.17	5.16	5.15	5.15	5.14	5.14	5.14	5.14	5.13	0.10	3
8.70	8.66	8.64	8.62	8.59	8.58	8.57	8.55	8.55	8.54	8.53	8.53	0.05	
26.9	26.7	26.6	26.5	26.4	26.4	26.3	26.2	26.2	26.2	26.1	26.1	0.01	
2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	0.25	
3.87	3.84	3.83	3.82	3.80	3.80	3.79	3.78	3.78	3.77	3.76	3.76	0.10	4
5.86	5.80	5.77	5.75	5.72	5.70	5.69	5.66	5.66	5.65	5.64	5.63	0.05	
14.2	14.0	13.9	13.8	13.7	13.7	13.7	13.6	13.6	13.5	13.5	13.5	0.01	
1.89	1.88	1.88	1.88	1.88	1.88	1.87	1.87	1.87	1.87	1.87	1.87	0.25	
3.24	3.21	3.19	3.17	3.16	3.15	3.14	3.13	3.12	3.12	3.11	3.10	0.10	5
4.62	4.56	4.53	4.50	4.46	4.44	4.43	4.41	4.40	4.39	4.37	4.36	0.05	
9.72	9.55	9.47	9.38	9.29	9.24	9.20	9.13	9.11	9.08	9.04	9.02	0.01	
1.76	1.76	1.75	1.75	1.75	1.75	1.74	1.74	1.74	1.74	1.74	1.74	0.25	
2.87	2.84	2.82	2.80	2.78	2.77	2.76	2.75	2.74	2.73	2.73	2.72	0.10	6
3.94	3.87	3.84	3.81	3.77	3.75	3.74	3.71	3.70	3.69	3.68	3.67	0.05	
7.56	7.40	7.31	7.23	7.14	7.09	7.06	6.99	6.97	6.93	6.90	6.88	0.01	
1.68	1.67	1.67	1.66	1.66	1.66	1.65	1.65	1.65	1.65	1.65	1.65	0.25	
2.63	2.59	2.58	2.56	2.54	2.52	2.51	2.50	2.49	2.48	2.48	2.47	0.10	7
3.51	3.44	3.41	3.38	3.34	3.32	3.30	3.27	3.27	3.25	3.24	3.23	0.05	
6.31	6.16	6.07	5.99	5.91	5.86	5.82	5.75	5.74	5.70	5.67	5.65	0.01	
1.62	1.61	1.60	1.60	1.59	1.59	1.59	1.58	1.58	1.58	1.58	1.58	0.25	
2.46	2.42	2.40	2.38	2.36	2.35	2.34	2.32	2.32	2.31	2.30	2.29	0.10	8
3.22	3.15	3.12	3.08	3.04	2.02	3.01	2.97	2.97	2.95	2.94	2.93	0.05	
5.52	5.36	5.28	5.20	5.12	5.07	5.03	4.96	4.95	4.91	4.88	4.86	0.01	
1.57	1.56	1.56	1.55	1.55	1.54	1.54	1.53	1.53	1.53	1.53	1.53	0.25	
2.34	2.30	2.28	2.25	2.23	2.22	2.21	2.19	2.18	2.17	2.17	2.16	0.10	9
3.01	2.94	2.90	2.86	2.83	2.80	2.79	2.76	2.75	2.73	2.72	2.71	0.05	
4.96	4.81	4.73	4.65	4.57	4.52	4.48	4.42	4.40	4.36	4.33	4.31	0.01	

分母N_2的自由度	Pr	分子N_1的自由度 1	2	3	4	5	6	7	8	9	10	11	12
	0.25	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.55	1.54
10	0.10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.30	2.28
	0.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
	0.01	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71
	0.25	1.47	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.53	1.52	1.52	1.51
11	0.10	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.23	2.21
	0.05	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
	0.01	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40
	0.25	1.46	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.51	1.50	1.50	1.49
12	0.10	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.17	2.15
	0.05	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
	0.01	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16
	0.25	1.45	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.47
13	0.10	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.12	2.10
	0.05	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60
	0.01	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96
	0.25	1.44	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.46	1.45
14	0.10	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.08	2.05
	0.05	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53
	0.01	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80
	0.25	1.43	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.46	1.45	1.44	1.44
15	0.10	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.04	2.02
	0.05	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48
	0.01	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67
	0.25	1.42	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.44	1.44	1.43
16	0.10	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	2.01	1.99
	0.05	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42
	0.01	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.62	3.55
	0.25	1.42	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.42	1.41
17	0.10	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.98	1.96
	0.05	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38
	0.01	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.46
	0.25	1.41	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42	1.41	1.40
18	0.10	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.96	1.93
	0.05	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
	0.01	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.43	3.37
	0.25	1.41	1.49	1.49	1.47	1.46	1.44	1.43	1.42	1.41	1.41	1.40	1.40
19	0.10	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.94	1.91
	0.05	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31
	0.01	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30
	0.25	1.40	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.41	1.40	1.39	1.39
20	0.10	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.92	1.89
	0.05	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28
	0.0l	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23

(续)

分子N_1的自由度													分母N_2的
15	20	24	30	40	50	60	100	120	200	500	∞	Pr	自由度
1.53	1.52	1.52	1.51	1.51	1.50	1.50	1.49	1.49	1.49	1.48	1.48	0.25	
2.24	2.20	2.18	2.16	2.13	2.12	2.11	2.09	2.08	2.07	2.06	2.06	0.10	10
2.85	2.77	2.74	2.70	2.66	2.64	2.62	2.59	2.58	2.56	2.55	2.54	0.05	
4.56	4.41	4.33	4.25	4.17	4.12	4.08	4.01	4.00	3.96	3.93	3.91	0.01	
1.50	1.49	1.49	1.48	1.47	1.47	1.47	1.46	1.46	1.46	1.45	1.45	0.25	
2.17	2.12	2.10	2.08	2.05	2.04	2.03	2.00	2.00	1.99	1.98	1.97	0.10	11
2.72	2.65	2.61	2.57	2.53	2.51	2.49	2.46	2.45	2.43	2.42	2.40	0.05	
4.25	4.10	4.02	3.94	3.86	3.81	3.78	3.71	3.69	3.66	3.62	3.60	0.01	
1.48	1.47	1.46	1.45	1.45	1.44	1.44	1.43	1.43	1.43	1.42	1.42	0.25	
2.10	2.06	2.04	2.01	1.99	1.97	1.96	1.94	1.93	1.92	1.91	1.90	0.10	12
2.62	2.54	2.51	2.47	2.43	2.40	2.38	2.35	2.34	2.32	2.31	2.30	0.05	
4.01	3.86	3.78	3.70	3.62	3.57	3.54	3.47	3.45	3.41	3.38	3.36	0.01	
1.46	1.45	1.44	1.43	1.42	1.42	1.42	1.41	1.41	1.40	1.40	1.40	0.25	
2.05	2.01	1.98	1.96	1.93	1.92	1.90	1.88	1.88	1.86	1.85	1.85	0.10	13
2.53	2.46	2.42	2.38	2.34	2.31	2.30	2.26	2.25	2.23	2.22	2.21	0.05	
3.82	3.66	3.59	3.51	3.43	3.38	3.34	3.27	3.25	3.22	3.19	3.17	0.01	
1.44	1.43	1.42	1.41	1.41	1.40	1.40	1.39	1.39	1.39	1.38	1.38	0.25	
2.01	1.96	1.94	1.91	1.89	1.87	1.86	1.83	1.83	1.82	1.80	1.80	0.10	14
2.46	2.39	2.35	2.31	2.27	2.24	2.22	2.19	2.18	2.16	2.14	2.13	0.05	
3.66	3.51	3.43	3.35	3.27	3.22	3.18	3.11	3.09	3.06	3.03	3.00	0.01	
1.43	1.41	1.41	1.40	1.39	1.39	1.38	1.38	1.37	1.37	1.36	1.36	0.25	
1.97	1.92	1.90	1.87	1.85	1.83	1.82	1.79	1.79	1.77	1.76	1.76	0.10	15
2.40	2.33	2.29	2.25	2.20	2.18	2.16	2.12	2.11	2.10	2.08	2.07	0.05	
3.52	3.37	3.29	3.21	3.13	3.08	3.05	2.98	2.96	2.92	2.89	2.87	0.01	
1.41	1.40	1.39	1.38	1.37	1.37	1.36	1.36	1.35	1.35	1.34	1.34	0.25	
1.94	1.89	1.87	1.84	1.81	1.79	1.78	1.76	1.75	1.74	1.73	1.72	0.10	16
2.35	2.28	2.24	2.19	2.15	2.12	2.11	2.07	2.06	2.04	2.02	2.01	0.05	
3.41	3.26	3.18	3.10	3.02	2.97	2.93	2.86	2.84	2.81	2.78	2.75	0.01	
1.40	1.39	1.38	1.37	1.36	1.35	1.35	1.34	1.34	1.34	1.33	1.33	0.25	
1.91	1.86	1.84	1.81	1.78	1.76	1.75	1.73	1.72	1.71	1.69	1.69	0.10	17
2.31	2.23	2.19	2.15	2.10	2.08	2.06	2.02	2.01	1.99	1.97	1.96	0.05	
3.31	3.16	3.08	3.00	2.92	2.87	2.83	2.76	2.75	2.71	2.68	2.65	0.01	
1.39	1.38	1.37	1.36	1.35	1.34	1.34	1.33	1.33	1.32	1.32	1.32	0.25	
1.89	1.84	1.81	1.78	1.75	1.74	1.72	1.70	1.69	1.68	1.67	1.66	0.10	18
2.27	2.19	2.15	2.11	2.06	2.04	2.02	1.98	1.97	1.95	1.93	1.92	0.05	
3.23	3.08	3.00	2.92	2.84	2.78	2.75	2.68	2.66	2.62	2.59	2.57	0.01	
1.38	1.37	1.36	1.35	1.34	1.33	1.33	1.32	1.32	1.31	1.31	1.30	0.25	
1.86	1.81	1.79	1.76	1.73	1.71	1.70	1.67	1.67	1.65	1.64	1.63	0.10	19
2.23	2.16	2.11	2.07	2.03	2.00	1.98	1.94	1.93	1.91	1.89	1.88	0.05	
3.15	3.00	2.92	2.84	2.76	2.71	2.67	2.60	2.58	2.55	2.51	2.49	0.01	
1.37	1.36	1.35	1.34	1.33	1.33	1.32	1.31	1.31	1.30	1.30	1.29	0.25	
1.84	1.79	1.77	1.74	1.71	1.69	1.68	1.65	1.64	1.63	1.62	1.61	0.10	20
2.20	2.12	2.08	2.04	1.99	1.97	1.95	1.91	1.90	1.88	1.86	1.84	0.05	
3.09	2.94	2.86	2.78	2.69	2.64	2.61	2.54	2.52	2.48	2.44	2.42	0.01	

分母N_2的自由度	Pr	分子N_1的自由度 1	2	3	4	5	6	7	8	9	10	11	12
	0.25	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.38	1.37
22	0.10	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.88	1.86
	0.05	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23
	0.01	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
	0.25	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.37	1.36
24	0.10	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.85	1.83
	0.05	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18
	0.01	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03
	0.25	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.36	1.35
26	0.10	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.84	1.81
	0.05	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	0.01	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96
	0.25	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
28	0.10	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.81	1.79
	0.05	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12
	0.01	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90
	0.25	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.35	1.34
30	0.10	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77
	0.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
	0.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84
	0.25	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31
40	0.10	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.73	1.71
	0.05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00
	0.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66
	0.25	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.29
60	0.10	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66
	0.05	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	0.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
	0.25	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.27	1.26
120	0.10	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.62	1.60
	0.05	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83
	0.01	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34
	0.25	1.33	1.39	1.38	1.36	1.34	1.32	1.31	1.29	1.28	1.27	1.26	1.25
200	0.10	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.60	1.57
	0.05	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80
	0.01	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27
	0.25	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.24
∞	0.10	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.57	1.55
	0.05	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75
	0.01	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18

资料来源： E.S. Pearson and H.O.Hartley,eds, *Biometrika Tables for Statisticians*, vol.1, 3d ed., table 18,Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.

(续)

分子N_1的自由度													分母N_2的
15	20	24	30	40	50	60	100	120	200	500	∞	Pr	自由度
1.36	1.34	1.33	1.32	1.31	1.31	1.30	1.30	1.30	1.29	1.29	1.28	0.25	
1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.61	1.60	1.59	1.58	1.57	0.10	22
2.15	2.07	2.03	1.98	1.94	1.91	1.89	1.85	1.84	1.82	1.80	1.78	0.05	
2.98	2.83	2.75	2.67	2.58	2.53	2.50	2.42	2.40	2.36	2.33	2.31	0.01	
1.35	1.33	1.32	1.31	1.30	1.29	1.29	1.28	1.28	1.27	1.27	1.26	0.25	
1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.58	1.57	1.56	1.54	1.53	0.10	24
2.11	2.03	1.98	1.94	1.89	1.86	1.84	1.80	1.79	1.77	1.75	1.73	0.05	
2.89	2.74	2.66	2.58	2.49	2.44	2.40	2.33	2.31	2.27	2.24	2.21	0.01	
1.34	1.32	1.31	1.30	1.29	1.28	1.28	1.26	1.26	1.26	1.25	1.25	0.25	
1.76	1.71	1.68	1.65	1.61	1.59	1.58	1.55	1.54	1.53	1.51	1.50	0.10	26
2.07	1.99	1.95	1.90	1.85	1.82	1.80	1.76	1.75	1.73	1.71	1.69	0.05	
2.81	2.66	2.58	2.50	2.42	2.36	2.33	2.25	2.23	2.19	2.16	2.13	0.01	
1.33	1.31	1.30	1.29	1.28	1.27	1.27	1.26	1.25	1.25	1.24	1.24	0.25	
1.74	1.69	1.66	1.63	1.59	1.57	1.56	1.53	1.52	1.50	1.49	1.48	0.10	28
2.04	1.96	1.91	1.87	1.82	1.79	1.77	1.73	1.71	1.69	1.67	1.65	0.05	
2.75	2.60	2.52	2.44	2.35	2.30	2.26	2.19	2.17	2.13	2.09	2.06	0.01	
1.32	1.30	1.29	1.28	1.27	1.26	1.26	1.25	1.24	1.24	1.23	1.23	0.25	
1.72	1.67	1.64	1.61	1.57	1.55	1.54	1.51	1.50	1.48	1.47	1.46	0.10	30
2.01	1.93	1.89	1.84	1.79	1.76	1.74	1.70	1.68	1.66	1.64	1.62	0.05	
2.70	2.55	2.47	2.39	2.30	2.25	2.21	2.13	2.11	2.07	2.03	2.01	0.01	
1.30	1.28	1.26	1.25	1.24	1.23	1.22	1.21	1.21	1.20	1.19	1.19	0.25	
1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.43	1.42	1.41	1.39	1.38	0.10	40
1.92	1.84	1.79	1.74	1.69	1.66	1.64	1.59	1.58	1.55	1.53	1.51	0.05	
2.52	2.37	2.29	2.20	2.11	2.06	2.02	1.94	1.92	1.87	1.83	1.80	0.01	
1.27	1.25	1.24	1.22	1.21	1.20	1.19	1.17	1.17	1.16	1.15	1.15	0.25	
1.60	1.54	1.51	1.48	1.44	1.41	1.40	1.36	1.35	1.33	1.31	1.29	0.10	60
1.84	1.75	1.70	1.65	1.59	1.56	1.53	1.48	1.47	1.44	1.41	1.39	0.05	
2.35	2.20	2.12	2.03	1.94	1.88	1.84	1.75	1.73	1.68	1.63	1.60	0.01	
1.24	1.22	1.21	1.19	1.18	1.17	1.16	1.14	1.13	1.12	1.11	1.10	0.25	
1.55	1.48	1.45	1.41	1.37	1.34	1.32	1.27	1.26	1.24	1.21	1.19	0.10	120
1.75	1.66	1.61	1.55	1.50	1.46	1.43	1.37	1.35	1.32	1.28	1.25	0.05	
2.19	2.03	1.95	1.86	1.76	1.70	1.66	1.56	1.53	1.48	1.42	1.38	0.01	
1.23	1.21	1.20	1.18	1.16	1.14	1.12	1.11	1.10	1.09	1.08	1.06	0.25	
1.52	1.46	1.42	1.38	1.34	1.31	1.28	1.24	1.22	1.20	1.17	1.14	0.10	200
1.72	1.62	1.57	1.52	1.46	1.41	1.39	1.32	1.29	1.26	1.22	1.19	0.05	
2.13	1.97	1.89	1.79	1.69	1.63	1.58	1.48	1.44	1.39	1.33	1.28	0.01	
1.22	1.19	1.18	1.16	1.14	1.13	1.12	1.09	1.08	1.07	1.04	1.00	0.25	
1.49	1.42	1.38	1.34	1.30	1.26	1.24	1.18	1.17	1.13	1.08	1.00	0.10	∞
1.67	1.57	1.52	1.46	1.39	1.35	1.32	1.24	1.22	1.17	1.11	1.00	0.05	
2.04	1.88	1.79	1.70	1.59	1.52	1.47	1.36	1.32	1.25	1.15	1.00	0.01	

表A-4 χ^2分布的

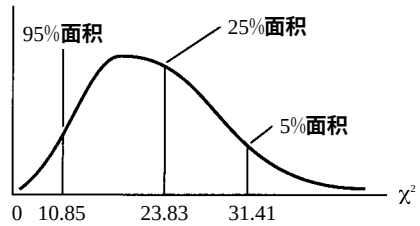
例:
Pr (χ^2＞10.85)=0.95
Pr (χ^2＞23.83)=0.25 自由度为20
Pr (χ^2＞31.41)=0.05

自由度 \ Pr	0.995	0.990	0.975	0.950	0.900
1	392 704×10^{-10}	157088×10^{-9}	982 069×10^{-9}	393 214×10^{-8}	0.0 157 908
2	0.0 100 251	0.0 201 007	0.0 506 356	0.102 587	0.210 720
3	0.0 717 212	0.114 832	0.2 157 95	0.351 846	0.584 375
4	0.2 06 990	0.297 110	0.484 419	0.710 721	1.063 623
5	0.411 740	0.554 300	0.831 211	1.145 476	1.61 031
6	0.675 727	0.872 085	1.237 347	1.63 539	2.20 413
7	0.989 265	1.239 043	1.68 987	2.16 735	2.83 311
8	1.344 419	1.646 482	2.17 973	2.73 264	3.48 954
9	1.734 926	2.087 912	2.70 039	3.32 511	4.16 816
10	2.15 585	2.55 821	3.24 697	3.94 030	4.86 518
11	2.60 321	3.05 347	3.81 575	4.57 481	5.57 779
12	3.07 382	3.57 056	4.40 379	5.22 603	6.30 380
13	3.56 503	4.10 691	5.00 874	5.89 186	7.04 150
14	4.07 468	4.66 043	5.62 872	6.57 063	7.78 953
15	4.60 094	5.22 935	6.26 214	7.26 094	8.54 675
16	5.14 224	5.81 221	6.90 766	7.96 164	9.31 223
17	5.69 724	6.40 776	7.56 418	8.67 176	10.0 852
18	6.26 481	7.01 491	8.23 075	9.39 046	10.8 649
19	6.84 398	7.63 273	8.90 655	10.1 170	11.6 509
20	7.43 386	8.26 040	9.59 083	10.8 508	12.4 426
21	8.03 366	8.89 720	10.28 293	11.5 913	13.2 396
22	8.64 272	9.54 249	10.9 823	12.3 380	14.0 415
23	9.26 042	10.19 567	11.6 885	13.0 905	14.8 479
24	9.88 623	10.8 564	12.4 011	13.8 484	15.6 587
25	10.5 197	11.5 240	13.1 197	14.6 114	16.4 734
26	11.1 603	12.1 981	13.8 439	15.3 791	17.2 919
27	11.8 076	12.8 786	14.5 733	16.1 513	18.1 138
28	12.4 613	13.5 648	15.3 079	16.9 279	18.9 392
29	13.1 211	14.2 565	16.0 471	17.7 083	19.7 677
30	13.7 867	14.9 535	16.7 908	18.4 926	20.5 992
40	20.7 065	22.1 643	24.4 331	26.5 093	29.0 505
50	27.9 907	29.7 067	32.3 574	34.7 642	37.6 886
60	35.5 346	37.4 848	40.4 817	43.1 879	46.4 589
70	43.2 752	45.4 418	48.7 576	51.7 393	55.3 290
80	51.1 720	53.5 400	57.1532	60.3 915	64.2 778
90	59.1 963	61.7 541	65.6 466	69.1 260	73.2 912
100①	67.3 276	70.0 648	74.2 219	77.9 295	82.3 581

① 当自由度大于100时，$\sqrt{2\chi^2}-\sqrt{(2k-1)}=Z$服从标准正态分布，其中，$k$代表自由度。

资料来源：E.S. Pearson and H.O.Hartley,eds, *Biometrika Tables for Statisticians*, vol.1, 3ded.,table 8,Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.

上百分点



0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
0.1 015 308	0.454 937	1.32 330	2.70 554	3.84 146	5.02 389	6.63 490	7.87 944
0.575 364	1.38 629	2.77 259	4.60 517	5.99 147	7.37 776	9.21 034	10.5 966
1.212 534	2.36 597	4.10 835	6.25 139	7.81 473	9.34 840	11.3 449	12.8 381
1.92 255	3.35 670	5.38 527	7.77 944	9.48 773	11.1 433	13.2 767	14.8 602
2.67 460	4.35 146	6.62 568	9.23 635	11.0 705	12.8 325	15.0 863	16.7 496
3.45 460	5.34 812	7.84 080	10.6 446	12.5 916	14.4 494	16.8 1 19	18.5 476
4.25 485	6.34 581	9.03 715	12.0 170	14.0 671	16.0 128	18.4 753	20.2 777
5.07 064	7.34 412	10.2 188	13.3 616	15.5 073	17.5 346	20.0 902	21.9 550
5.89 883	8.34 283	11.3 887	14.6 837	16.9 190	19.0 228	21.6 660	23.5 893
6.73 720	9.34 182	12.5 489	15.9 871	18.3 070	20.4 831	23.2 093	25.1 882
7.58 412	10.3 410	13.7 007	17.2 750	19.6 751	21.9 200	24.7 250	26.7 569
8.43 842	11.3 403	14.8 454	18.5 494	21.0 261	23.3 367	26.2 170	28.2 995
9.29 906	12.3 398	15.9 839	19.8 1 19	22.3 621	24.7 356	27.6 883	29.8 194
10.1 653	13.3 393	17.1 170	21.0 642	23.6 848	26.1 190	29.1 413	31.3 193
11.0 365	14.3 389	18.2 451	22.3 072	24.9 958	27.4 884	30.5 779	32.8 013
11.9 122	15.3 385	19.3 688	23.5 418	26.2 962	28.8 454	31.9 999	34.2 672
12.7 919	16.3 381	20.4 887	24.7 690	27.5 871	30.1 910	33.4 087	35.7 185
13.6 753	17.3 379	21.6 049	25.9 894	28.8 693	31.5 264	34.8 053	37.1 564
14.5 620	18.3 376	22.7 178	27.2 036	30.1 435	32.8 523	36.1 908	38.5 822
15.4 518	19.3 374	23.8 277	28.4 120	31.4 104	34.1 696	37.5 662	39.9 968
16.3 444	20.3 372	24.9 348	29.6 151	32.6 705	35.4 789	38.9 321	41.4 010
17.2 396	21.3 370	26.0 393	30.8 133	33.9 244	36.7 807	40.2 894	42.7 956
18.1 373	22.3 369	27.1 413	32.0 069	35.1 725	38.0 757	41.6 384	44.1 813
19.0 372	23.3367	28.2 412	33.1 963	36.4 151	39.3 641	42.9 798	45.5 585
19.9 393	24.3 366	29.3 389	34.3 816	37.6 525	40.6 465	44.3 141	46.9 278
20.8 434	25.3 364	30.4 345	35.5 631	38.8 852	41.9 232	45.6 417	48.2 899
21.7 494	26.3 363	31.5 284	36.7 412	40.1 133	43.1 944	46.9 630	49.6 449
22.6 572	27.3 363	32.6 205	37.9 159	41.3 372	44.4 607	48.2 782	50.9 933
23.5 666	28.3 362	33.7 109	39.0 875	42.5 569	45.7 222	49.5 879	52.3 356
24.4 776	29.3 360	34.7 998	40.2 560	43.7 729	46.9 792	50.8 922	53.6 720
33.6 603	39.3 354	45.6 160	51.8 050	55.7 585	59.3 417	63.6 907	66.7 659
42.9 421	49.3 349	56.3 336	63,1 671	67.5 048	71.4 202	76.1 539	79.4 900
52.2 938	59.3 347	66.9 814	74.3 970	79.0 819	83.2 976	88.3 794	91.9 517
61.6 983	69.3 344	77.5 766	85.5 271	90.5 312	95.0 231	100.425	104. 215
71.1 445	79.3 343	88.1 303	96.5 782	101.879	106.629	112.329	116.321
80.6 247	89.3 342	98.6 499	107.565	113.145	118.136	124.116	128.299
90.1 332	99.3 341	109.141	118.498	124.342	129.561	135.807	140.169

表A-5a 杜宾-瓦尔森d统计量：显著水平0.05的d_L与d_U的显著点

	k'=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7		k'=8		k'=9		k'=10	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.610	1.400	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.368	2.287	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.824	1.320	0.629	1,699	0.455	2.128	0.296	2.588	—	—	—	—	—	—	—	—	—	—	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	—	—	—	—	—	—	—	—	—	—
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005	—	—	—	—	—	—	—	—
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149	—	—	—	—	—	—
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266	—	—	—	—
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	—	—
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	1.239	1.429	1.147	1.541	1.053	1,66A	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	2.174	0.666	2.318	0.584	2.464	0.506	2.613
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470
28	1.328	1.476	1.255	1.560	1.181	1,650	1.104	1.747	1.028	1.850	0.951	1.958	0.874	2.071	0.798	2.188	0.723	2.309	0.650	2.431
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.682	2.396
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.795	2.281
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.080	1.891	1.015	1.979	0.950	2.069	0.885	2.162	0.821	2.257
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.877	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.198
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.952	2.149
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1,837	1.369	1.873	1.337	1.910	1.305	1.948
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916
90	1.635	1.679	1.612	1,703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898
150	1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.847	1.608	1.862	1.594	1.877
200	1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874

(续)

	k'=11		k'=12		k'=13		k'=14		k'=15		k'=16		k'=17		k'=18		k'=19		k'=20	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
16	0.098	3.503	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0.138	3.378	0.087	3.557	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0.177	3.265	0.123	3.441	0.078	3.603	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0.220	3.159	0.160	3.350	0.111	3.496	0.070	3.642	—	—	—	—	—	—	—	—	—	—	—	—
20	0.263	3.063	0.200	3.234	0.145	3.395	0.100	3.542	0.063	3.676	—	—	—	—	—	—	—	—	—	—
21	0.307	2.976	0.240	3.141	0.182	3.300	0.132	3.448	0.091	3.583	0.058	3.705	—	—	—	—	—	—	—	—
22	0.349	2.897	0.281	3.057	0.220	3.211	0.166	3.358	0.120	3.495	0.083	3.619	0.052	3.731	—	—	—	—	—	—
23	0.391	2.826	0.322	2.979	0.259	3.128	0.202	3.272	0.153	3.409	0.110	3.535	0.076	3.650	0.048	3.753	—	—	—	—
24	0.431	2.761	0.362	2.908	0.297	3.053	0.239	3.193	0.186	3.327	0.141	3.454	0.101	3.572	0.070	3.678	0.044	3.773	—	—
25	0.470	2.702	0.400	2.844	0.335	2.983	0.275	3.119	0.221	3.251	0.172	3.376	0.130	3.494	0.094	3.604	0.065	3.702	0.041	3.790
26	0.508	2.649	0.438	2.784	0.373	2.919	0.312	3.051	0.256	3.179	0.205	3.303	0.160	3.420	0.120	3.531	0.087	3.632	0.060	3.724
27	0.544	2.600	0.475	2.730	0.409	2.859	0.348	2.987	0.291	3.112	0.238	3.233	0.191	3.349	0.149	3.460	0.112	3.563	0.081	3.658
28	0.578	2.555	0.510	2.680	0.445	2.805	0.383	2.928	0.325	3.050	0.271	3.168	0.222	3.283	0.178	3.392	0.138	3.495	0.104	3.592
29	0.612	2.515	0.544	2.634	0.479	2.755	0.418	2.874	0.359	2.992	0.305	3.107	0.254	3.219	0.208	3.327	0.166	3.431	0.129	3.528
30	0.643	2.477	0.577	2.592	0.512	2.708	0.451	2.823	0.392	2.937	0.337	3.050	0.286	3.160	0.238	3.266	0.195	3.368	0.156	3.465
31	0.674	2.443	0.608	2.553	0.545	2.665	0.484	2.776	0.425	2.887	0.370	2.996	0.317	3.103	0.269	3.208	0.224	3.309	0.183	3.406
32	0.703	2.411	0.638	2.517	0.576	2.625	0.515	2.733	0.457	2.840	0.401	2.946	0.349	3.050	0.299	3.153	0.253	3.252	0.211	3.348
33	0.731	2.382	0.668	2.484	0.606	2.588	0.546	2.692	0.488	2.796	0.432	2.899	0.379	3.000	0.329	3.100	0.283	3.198	0.239	3.293
34	0.758	2.355	0.695	2.454	0.634	2.554	0.575	2.654	0.518	2.754	0.462	2.854	0.409	2.954	0.359	3.051	0.312	3.147	0.267	3.240
35	0.783	2.330	0.722	2.425	0.662	2.521	0.604	2.619	0.547	2.716	0.492	2.813	0.439	2.910	0.388	3.005	0.340	3.099	0.295	3.190
36	0.808	2.306	0.748	2.398	0.689	2.492	0.631	2.586	0.575	2.680	0.520	2.774	0.467	2.868	0.417	2.961	0.369	3.053	0.323	3.142
37	0.831	2.285	0.772	2.374	0.714	2.464	0.657	2.555	0.602	2.646	0.548	2.738	0.495	2.829	0.445	2.920	0.397	3.009	0.351	3.097
38	0.854	2.265	0.796	2.351	0.739	2.438	0.683	2.526	0.628	2.614	0.575	2.703	0.522	2.792	0.472	2.880	0.424	2.968	0.378	3.054
39	0.875	2.246	0.819	2.329	0.763	2.413	0.707	2.499	0.653	2.585	0.600	2.671	0.549	2.757	0.499	2.843	0.451	2.929	0.404	3.013
40	0.896	2.228	0.840	2.309	0.785	2.391	0.731	2.473	0.678	2.557	0.626	2.641	0.575	2.724	0.525	2.808	0.477	2.892	0.430	2.974
45	0.988	2.156	0.938	2.225	0.887	2.296	0.838	2.367	0.788	2.439	0.740	2.512	0.692	2.586	0.644	2.659	0.598	2.733	0.553	2.807
50	1.064	2.103	1.019	2.163	0.973	2.225	0.927	2.287	0.882	2.350	0.836	2.414	0.792	2.479	0.747	2.544	0.703	2.610	0.660	2.675
55	1.129	2.062	1.087	2.116	1.045	2.170	1.003	2.225	0.961	2.281	0.919	2.338	0.877	2.396	0.836	2.454	0.795	2.512	0.754	2.571
60	1.184	2.031	1.145	2.079	1.106	2.127	1.068	2.177	1.029	2.227	0.990	2.278	0.951	2.330	0.913	2.382	0.874	2.434	0.836	2.487
65	1.231	2.006	1.195	2.049	1.160	2.093	1.124	2.138	1.088	2.183	1.052	2.229	1.016	2.276	0.980	2.323	0.944	2.371	0.908	2.419
70	1.272	1.986	1.239	2.026	1.206	2.066	1.172	2.106	1.139	2.148	1.105	2.189	1.072	2.232	1.038	2.275	1.005	2.318	0.971	2.362
75	1.308	1.970	1.277	2.006	1.247	2.043	1.215	2.080	1.184	2.118	1,153	2.156	1.121	2.195	1.090	2.235	1.058	2.275	1.027	2.315
80	1.340	1.957	1.311	1.991	1.283	2.024	1.253	2.059	1.224	2.093	1.195	2.129	1.165	2.165	1.136	2.201	1.106	2.238	1.076	2.275
85	1.369	1.946	1.342	1.977	1.315	2.009	1.287	2.040	1.260	2.073	1.232	2.105	1.205	2.139	1.177	2.172	1.149	2.206	1.121	2.241
90	1.395	1.937	1.369	1.966	1.344	1.995	1.318	2.025	1.292	2.055	1.266	2.085	1.240	2.116	1.213	2.148	1.187	2.179	1.160	2.211
95	1.418	1.929	1.394	1.956	1.370	1.984	1.345	2.012	1.321	2.040	1.296	2.068	1.271	2.097	1.247	2.126	1.222	2.156	1.197	2.186
100	1.439	1.923	1.416	1.948	1.393	1.974	1.371	2.000	1.347	2.026	1.324	2.053	1.301	2.080	1.277	2.108	1.253	2.135	1.229	2.164
150	1.579	1.892	1.564	1.908	1.550	1.924	1.535	1.940	1.519	1.956	1.504	1.972	1.489	1.989	1.474	2.006	1.458	2.023	1.443	2.040
200	1.654	1.885	1.643	1.896	1.632	1.908	1.621	1.919	1.610	1.931	1.599	1.943	1.5588	1.955	1.576	1.967	1.565	1.979	1.554	1.991

资料来源： This table is an extension of the original Durbin-Watson table and is reproduced from N.E.Savin and K.J.White, "The Dubin-Watson Test for Serial Correlation with Extreme Small Samples or Many Regressors," *Econometrica*,vol.45,November 1977,pp.1989-96 and as corrected by R.W.Farebrother, *Econometrica* ,vol.48,September 1980,p.1554. Reprinted by permission of the Econometric Society.

例：若n=40, k'=4, d_L=1.285, d_U=1.721。如果计算的d值小于1.85，则表明存在一阶序列相关；如果大于1.721，则没有证据表明存在一阶序列相关；但是，如果 d位于上、下限之间，则无法确定是否存在一阶序列相关。

表A-5b 杜宾-瓦尔森d统计量：显著水平0.01的d_L与d_U的显著点

	k'=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7		k'=8		k'=9		k'=10	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.390	1.142	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.435	1.036	0.294	1.676	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.497	1.003	0.345	1.489	0.229	2.102	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.554	0.998	0.408	1.389	0.279	1.875	0.183	2.433	—	—	—	—	—	—	—	—	—	—	—	—
10	0.604	1.001	0.466	1.333	0.340	1.733	0.230	2.193	0.150	2.690	—	—	—	—	—	—	—	—	—	—
11	0.653	1.010	0.519	1.297	0.396	1.640	0.286	2.030	0.193	2.453	0.124	2.892	—	—	—	—	—	—	—	—
12	0.697	1.023	0.569	1.274	0.449	1.575	0.339	1.913	0.244	2.280	0.164	2.665	0.105	3.053	—	—	—	—	—	—
13	0.738	1.038	0.616	1.261	0.499	1.526	0.391	1.826	0.294	2.150	0.211	2.490	0.140	2.838	0.090	3.182	—	—	—	—
14	0.776	1.054	0.660	1.254	0.547	1.490	0.441	1.757	0.343	2.049	0.257	2,354	0.183	2.667	0.122	2.981	0.078	3.287	—	—
15	0.811	1.070	0.700	1.252	0.591	1.464	0.488	1.704	0.391	1.967	0.303	2.244	0.226	2.530	0.161	2.817	0.107	3.101	0.068	3.374
16	0.844	1.086	0.737	1.252	0.633	1.446	0.532	1.663	0.437	1.900	0.349	2.153	0.269	2.416	0.200	2.681	0.142	2.944	0.094	3.201
17	0.874	1.102	0.772	1.255	0.672	1.432	0.574	1.630	0.480	1.847	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127	3.053
18	0.902	1.118	0.805	1.259	0.708	1.422	0.613	1.604	0.522	1.803	0.435	2.015	0.355	2.238	0.282	2.467	0.216	2.697	0.160	2.925
19	0.928	1.132	0.835	1.265	0.742	1.415	0.650	1.584	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.381	0.255	2.597	0.196	2.813
20	0.952	1.147	0.863	1.271	0.773	1.411	0.685	1.567	0.598	1.737	0.515	1.918	0.436	2.110	0.362	2.308	0.294	2.510	0.232	2.714
21	0.975	1.161	0.890	1.277	0.803	1.408	0.718	1.554	0.633	1.712	0.552	1.881	0.474	2.059	0.400	2.244	0.331	2.434	0.268	2.625
22	0.997	1.174	0.914	1.284	0.831	1.407	0.748	1.543	0.667	1.691	0.587	1.849	0.510	2.015	0.437	2.188	0.368	2.367	0.304	2.548
23	1.018	1.187	0.938	1.291	0.858	1.407	0.777	1.534	0.698	1.673	0.620	1.821	0.545	1.977	0.473	2.140	0.404	2.308	0.304	2.479
24	1.037	1.199	0.960	1.298	0.882	1.407	0.805	1.528	0.728	1.658	0.652	1.797	0.578	1.944	0.507	2.097	0.439	2.255	0.375	2.417
25	1.055	1.211	0.981	1.305	0.906	1.409	0.831	1.523	0.756	1.645	0.682	1.776	0.610	1.915	0.540	2.059	0.473	2.209	0.409	2.362
26	1.072	1.222	1.001	1.312	0.928	1.411	0.855	1.518	0.783	1.635	0.711	1.759	0.640	1.889	0.572	2.026	0.505	2.168	0.441	2.313
27	1.089	1.233	1.019	1.319	0.949	1.413	0.878	1.515	0.808	1.626	0.738	1.743	0.669	1.867	0.602	1.997	0.536	2.131	0.473	2.269
28	1.104	1.244	1.037	1.325	0.969	1.415	0.900	1.513	0.832	1.618	0.764	1.729	0.696	1.847	0.630	1.970	0.566	2.098	0.504	2.229
29	1.119	1.254	1.054	1.332	0.988	1.418	0.921	1.512	0.855	1.611	0.788	1.718	0.723	1.830	0.658	1.947	0.595	2.068	0.533	2.193
30	1.133	1.263	1.070	1.339	1.006	1.421	0.941	1.511	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562	2.160
31	1.147	1.273	1.085	1.345	1.023	1.425	0.960	1.510	0.897	1.601	0.834	1.698	0.772	1.800	0.710	1.906	0.649	2.017	0.589	2.131
32	1.160	1.282	1.100	1.352	1.040	1.428	0.979	1.510	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615	2.104
33	1.172	1.291	1.114	1.358	1.055	1.432	0.996	1.510	0.936	1.594	0.876	1.683	0.816	1.776	0.757	1.874	0.698	1.975	0.641	2.080
34	1.184	1.299	1.128	1.364	1.070	1.435	1.012	1.511	0.954	1.591	0.896	1.677	0.837	1.766	0.779	1.860	0.722	1.957	0.665	2.057
35	1.195	1.307	1.140	1.370	1.085	1.439	1.028	1.512	0.971	1.589	0.914	1.671	0.857	1.757	0.800	1.847	0.744	1.940	0.689	2.037
36	1.206	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.988	1.588	0.932	1.666	0.877	1.749	0.821	1.836	0.766	1.925	0.711	2.018
37	1.217	1.323	1.165	1.382	1.112	1.446	1.058	1.514	1.004	1.586	0.950	1.662	0.895	1.742	0.841	1.825	0.787	1.911	0.733	2.001
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.585	0.966	1.658	0.913	1.735	0.860	1.816	0.807	1.899	0.754	1.985
39	1.237	1.337	1.187	1.393	1.137	1.453	1.085	1.517	1.034	1.584	0.982	1.655	0.930	1.729	0.878	1.807	0.826	1.887	0.774	1.970
40	1.246	1.344	1.198	1.398	1.148	1.457	1.098	1.518	1.048	1.584	0.997	1.652	0.946	1.724	0.895	1.799	0.844	1.876	0.749	1.956
45	1.288	1.376	1.245	1.423	1.201	1.474	1.156	1.528	1.111	1.584	1.065	1.643	1.019	1.704	0.974	1.768	0.927	1.834	0.881	1.902
50	1.324	1.403	1.285	1.446	1.245	1.491	1.205	1.538	1.164	1.587	1.123	1.639	1.081	1.692	1.039	1.748	0.997	1.805	0.955	1.864
55	1.356	1.427	1.320	1.466	1.284	1.506	1.247	1.548	1.209	1.592	1.172	1.638	1.134	1.685	1.095	1.734	1.057	1.785	1.018	1.837
60	1.383	1.449	1.350	1.484	1.317	1.520	1.283	1.558	1.249	1.598	1.214	1.639	1.179	1.682	1.144	1.726	1.108	1.771	1.072	1.817
65	1.407	1.468	1.377	1.500	1.346	1.534	1.315	1.568	1.283	1.604	1.251	1.642	1.218	1.680	1.186	1.720	1.153	1.761	1.120	1.802
70	1.429	1.485	1.400	1.515	1.372	1.546	1.343	1.578	1.313	1.611	1.283	1.645	1.253	1.680	1.223	1.716	1.192	1.754	1.162	1.792
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.587	1.340	1.617	1.313	1.649	1.284	1.682	1.256	1.714	1.227	1.748	1.199	1.783
80	1.466	1.515	1.441	1.541	1.416	1.568	1.390	1.595	1.364	1.624	1.338	1.653	1.312	1.683	1.285	1.714	1.259	1.745	1.232	1.777
85	1.482	1.528	1.458	1.553	1.435	1.578	1.411	1.603	1.386	1.630	1.362	1.657	1.337	1.685	1.312	1.714	1.287	1.743	1.262	1.773
90	1.496	1.540	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636	1.383	1.661	1.360	1.687	1.336	1.714	1.312	1.741	1.288	1.769
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.642	1.403	1.666	1.381	1.690	1.358	1.715	1.336	1.741	1.313	1.767
100	1.522	1.562	1.503	1.583	1.482	1.604	1.462	1.625	1.441	1.647	1.421	1.670	1.400	1.693	1.378	1.717	1.357	1.741	1.335	1.765
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693	1.543	1.708	1.530	1.722	1.515	1.737	1.501	1.752	1.486	1.767
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725	1.613	1.735	1.603	1.746	1.592	1.757	1.582	1.768	1.571	1.779

(续)

	k'=11		k'=12		k'=13		k'=14		k'=15		k'=16		k'=17		k'=18		k'=19		k'=20	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
16	0.060	3.446	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0.084	3.286	0.053	3.506	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0.13	3.146	0.075	3.358	0.047	3.357	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0.145	3.023	0.102	3.227	0.067	3.420	0.043	3.601	—	—	—	—	—	—	—	—	—	—	—	—
20	0.178	2.914	0.131	3.109	0.092	3.297	0.061	3.474	0.038	3.639	—	—	—	—	—	—	—	—	—	—
21	0.212	2.817	0.162	3.004	0.119	3.185	0.084	3.358	0.055	3.521	0.035	3.671	—	—	—	—	—	—	—	—
22	0.246	2.729	0.194	2.909	0.148	3.084	0.109	3.252	0.077	3.412	0.050	3.562	0.032	3.700	—	—	—	—	—	—
23	0.281	2.651	0.227	2.822	0.178	2.991	0.136	3.155	0.100	3.311	0.070	3.459	0.046	3.597	0.029	3.725	—	—	—	—
24	0.315	2.580	0.260	2.744	0.209	2.906	0.165	3.065	0.125	3.218	0.092	3.363	0.065	3.501	0.043	3.629	0.027	3.747	—	—
25	0.348	2.517	0.292	2.674	0.240	2.829	0.194	2.982	0.152	3.131	0.116	3.274	0.085	3.410	0.060	3-538	0.039	3.657	0.025	3.766
26	0.381	2.460	0.324	2.610	0.272	2.758	0.224	2.906	0.180	3.0-500	.141	3.191	0.107	3.325	0.079	3.452	0.055	3.572	0.036	3.682
27	0.413	2.409	0.356	2.552	0.303	2.694	0.253	2.836	0.208	2.976	0.167	3.113	0.131	3.245	0.100	3.371	0.073	3.490	0.051	3.602
28	0.444	2.363	0.387	2.499	0.333	2.635	0.283	2.772	0.237	2.907	0.194	3.040	0.156	3.169	0.122	3.294	0.093	3.412	0.068	3.524
29	0.474	2.321	0.417	2.451	0.363	2.582	0.313	2.713	0.266	2.843	0.222	2.972	0.182	3.098	0.146	3.220	0.114	3.338	0.087	3.450
30	0.503	2.283	0.447	2.407	0.393	2.533	0.342	2.659	0.294	2.785	0.249	2.909	0.208	3.032	0.171	3.152	0.137	3.267	0.107	3.379
31	0.531	2.248	0.475	2.367	0.422	2.487	0.371	2.609	0.322	2.730	0.277	2.851	0.234	2.970	0.196	3.087	0.160	3.201	0.128	3.311
32	0.558	2.216	0.503	2.330	0.450	2.446	0.399	2.563	0.350	2.680	0.304	2.797	0.261	2.912	0.221	3.026	0.184	3.137	0.151	3.246
33	0.585	2.187	0.530	2.296	0.477	2.408	0.426	2.520	0.377	2.633	0.331	2.746	0.287	2.858	0.246	2.969	0.209	3.078	0.174	3.184
34	0.610	2.160	0.556	2.266	0.503	2.373	0.452	2.481	0.404	2.590	0.357	2.699	0.313	2.808	0.272	2.915	0.233	3.022	0.197	3.126
35	0.634	2.136	0.581	2.237	0.529	2.340	0.478	2.444	0.430	2.550	0.383	2.655	0.339	2.761	0.297	2.865	0.257	2.969	0.221	3.071
36	0.658	2.113	0.605	2.210	0.554	2.310	0.504	2.410	0.455	2.512	0.409	2.614	0.364	2.717	0.322	2.818	0.282	2.919	0,244	3.019
37	0.680	2.092	0.628	2.186	0.578	2.282	0.528	2.379	0.480	2.477	0.434	2.576	0.389	2.675	0.347	2.774	0.306	2.872	0.268	2.969
38	0.702	2.073	0.651	2.164	0.601	2.256	0.552	2.350	0.504	2.445	0.458	2.540	0.414	2.637	0.371	2.733	0.330	2.828	0.291	2.923
39	0.723	2.055	0.673	2.143	0.623	2.232	0.575	2.323	0.528	2.414	0.482	2.507	0.438	2.600	0.395	2.694	0.354	2.787	0.315	2.879
40	0.744	2.039	0.694	2.123	0.645	2.210	0.597	2.297	0.551	2.386	0.505	2.476	0.461	2.566	0.418	2.657	0.377	2.748	0.338	2.838
45	0.835	1.972	0.790	2.044	0.744	2.118	0.700	2.193	0.655	2.269	0.612	2.346	0.570	2.424	0.528	2.503	0.488	2.582	0.448	2.661
50	0.913	1.925	0.871	1.987	0.829	2.051	0.787	2.116	0.746	2.182	0.705	2.250	0.665	2.318	0.625	2.387	0.586	2.456	0.548	2.526
55	0.979	1.891	0.940	1.945	0.902	2.002	0.863	2.059	0.825	2.117	0.786	2.176	0.748	2.237	0.711	2.298	0.674	2.359	0.637	2.421
60	1.037	1.865	1.001	1.914	0.96-5	1.964	0.929	2.015	0.893	2.067	0.857	2.120	0.822	2.173	0.786	2.227	0.751	2.283	0.716	2.338
65	1.087	1.845	1.053	1.889	1.020	1.934	0.986	1.980	0.953	2.027	0.919	2.075	0.886	2.123	0.852	2.172	0.819	2.221	0.786	2.272
70	1.131	1.831	1.099	1.870	1.068	1.911	1.037	1.953	1.005	1.995	0.974	2.038	0.943	2.082	0.911	2.127	0.880	2.172	0.849	2.217
75	1.170	1.819	1.141	1.856	1.111	1.893	1.082	1.931	1.052	1.970	1.023	2.009	0.993	2.049	0.964	2.090	0.934	2.131	0.905	2.172
80	1.205	1.810	1.177	1.844	1.150	1.878	1.122	1.913	1.094	1.949	1.066	1.984	1.039	2.022	1.011	2.059	0.983	2.097	0.955	2.135
85	1.236	1.803	1.210	1.834	1.184	1.866	1.158	1.898	1.132	1.931	1.106	1.965	1.080	1.999	1.053	2.033	1.027	2.068	1.000	2.104
90	1.264	1.798	1.240	1.827	1.215	1.856	1.191	1.886	1.166	1.917	1.141	1.948	1.116	1.979	1.091	2.012	1.066	2.044	1.041	2.077
95	1.290	1.793	1.267	1.821	1.244	1.848	1.221	1.876	1.197	1.905	1.174	1.934	1.150	1.963	1.126	1.993	1.102	2.023	1.079	2.054
100	1.314	1.790	1.292	1.816	1.270	1.841	1.248	1.868	1.225	1.895	1.203	1.922	1.181	1.949	1.158	1.977	1.136	2.006	1.113	2.034
150	1.473	1.783	1.458	1.799	1.444	1.814	1.429	1.830	1.414	1.847	1.400	1.863	1.385	1.880	1.370	1.897	1.355	1.913	1.340	1.931
200	1.561	1.791	1.550	1.801	1.539	1.813	1.528	1.824	1.518	1.836	1.507	1.847	1.495	1.860	1.484	1.871	1.474	1.883	1.462	1.896

资料来源： Savin and White,op.cit.,by permission of the Econometric Society.

注： n —— **观察值的个数**

k' —— **解释变量的个数(不包括常数项)**

表A-6 游程检验中的各游程的临界值

	N_2																		
N_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	to	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

	N_2																		
N_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

注：表A-6a**与表**A-6b**对**N_1(+**号)与**N_2(-**号)的不同值，给出游程的临界值。一个样本的游程检验，显著水平为**0.05，*n*的任一值等于或小于表*D*所示临界值，或者*n*的任一值等于或大于表D 所示的临界值，则为显著。

资料来源：Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, New York,1956, table F,pp.252-253.The tables have been adapted by Siegel from the original source : Frieda S.Swed and C.Eisenhart, “Tables for Testing Randomness of Grouping in a Sequence of Alternatives,” *Annals of Mathematical Statistics*, vol.14,1943.Used by permission of McGraw-Hill Book Company and *Annals of Mathematical Statistics*.

例：在由30**个观察值组成的序列中，有** 20**个正号(**N_1=20)，10**个负号(**N_2=10)。在5%**的显著水平下，根据表** A-6a**和表**A-6b**知游程的临界值为** 9**和**20**。因此，若游程的个数小于或等于** 9**或是大于或等于**20**，则可以拒绝假设(在**5%**的显著水平下)：所观察的序列是随机的。**

附 录 B

■ 部分习题答案

2.14 (a) b=1/15;(b) 6/15,10/15,4/15; (c) 40/15; (d) 1.5556

2.18 (a) $E(X)$=13%; (b) $E(Y)$=34.7%; (c) $E(XY)$=635。不是独立的，因为E (X) E (Y)=451.1, E (XY)=635

2.23.0.558

2.24 (a) 均值=617 710,方差=$(65\ 499)^2$

(b) 均值=38 201,方差=$(21\ 935)^2$

(c) 协方差=0.139 93E+10,相关系数=0.973 97

(d) 不独立

(e) 不，因为相关并不意味着存在因果关系。

3.10 (a) Z=－4.17，获此Z值的概率非常小。

(b) 概率很小。

3.11 利润约为1.653 6百万美元。

3.13 (a)约为264; (b)52.8;(c)13.2

3.14 (a) N(25,11); (b) N (-5,11); (c) N (30,27); (d) N(115,248)

3.20 S_1^2/S_2^2服从自由度为(n_1-1)和(n_2-2)的F分布。F=1.25, 获此F值的概率为0.24。

4.7 (a)=1.96; (b)=1.65; (c)2.58; (d)=2.33

4.9 (a) 0.9544; (b)0.0228; (c)0.0228; (d)是

4.11 (a) μ=6.668 6; σ=0.102 5; (b) 约为0.000 612

4.14 (a) $\bar{X}$ ~N (8,36/25)

(b) Z=-0.416 7。获此Z值的概率约为0.34。

(c) 是,因为 $\bar{X}$ 的95%的置信区间为 (5.648,10.352)。样本均值为 7.5,位于该置信区间内。

4.19 (a)95%的置信区间为(9.253 5,34.132 3)

(b) 因为上述区间不包括8.2,所以拒绝假设。

5.6 (a) 是; (b)是; (c)是; (d)是; (e)不是; (f)不是

5.10 (a) 和(b): 散点图表明呈现正相关。

(c) $(S\&P)_t$=－195.08+3.8211CPI_t

se=(61.106) (0.609 7) r^2=0.797 1

(d) 与预期相同,正相关。

5.17 ASP_i=-267 875+103 423GPA_i

se=(86 042) (26 434)

t=(−3.110) (3.91) r^2=0.353

因为估计的GPA的系数是统计显著的,所以它对ASP有正的影响。

(b) ASP_i=-335 983+648 GMAT

se=(45 702) (73.16)

t=(−7.35) (8.86) r^2=0.737

显著正相关。

(c) ASP_i=22 504+2.69$Tuition_i$

se=(9 542) (0.5 318)

t=(2.36) (5.00)

显著正相关。

6.9 (a) b_1=21.22, b_2=0.5344

(b) se(b_1)=8.5913, se (b_2)=0.048 4

(c) r^2=0.935 8

(d) B_1=(1.408, 41.032)

B_2=(0.422 7, 0.646 0)

(e) **拒绝零假设。**

6.13 (a)**正。**

(b) **和**(c):**散点图表明两者之间是正相关。**

(d) $\hat{Y}_t$ =373.301 4+0.419 9X_t

se=(9 530.376 8) (0.115 4) r^2=0.546

(e) 0.061 6≤B_2≤0.778 3 **因为该置信区间不包括零假设值，所以拒绝零假设。**

6.18 (a) $\hat{Y}_t$=148.134 8+0.672 7 X_t

t=(12.712 0) (25.101 8) r^2=0.966 3

(b) **拒绝零假设，因为**t**值相当高。**

(c) $\hat{Y}_{1991}$ =434 (**近似值**)

(d) (432.915, 435.085)

7.8 (a), (b), (c):

$\hat{Y}_t$=53.159 9+0.726 6X_{2i}+2.7363X_{3i}

se= (13.026 1) (0.0487) (0.8454) R^2=0.998 8

t=(4.0429) (14.9197) (3.2367) $\bar{R}^2$ =0.9986

(d) B_2: (0.620 5,0.832 7)

B_3: (0.894 2,4.578 4)

(e) **根据上面的置信区间,两个斜率系数各自均是统计不独立的。**

(f) **根据**R^2**的值,估计的**F**值**4 994**是高度显著的。**

7.9 (a) 15; (b)77; (c)2**和**12; (d) R^2=0.9988, $\bar{R}^2$ =0.9986; (e) F=5 140.13, **高度显著的**, **因而拒绝零假设**; (f)**不**, **需要双变量回归模型的结果。**

7.10 **对**k**个变量的模型(包括截距项)，我们有**,

$$F=\frac{\mathrm{ESS}/(k-1)}{\mathrm{RSS}/(n-k)}=\frac{\mathrm{ESS}(n-k)}{\mathrm{RSS}(k-1)}=\frac{\mathrm{ESS}(n-k)}{(\mathrm{TSS}-\mathrm{ESS})(k-1)}=\frac{\mathrm{ESS}/\mathrm{TSS}(n-k)}{[1-(\mathrm{ESS}/\mathrm{TSS})](k-1)}$$

$$=\frac{R^2/(k-1)}{(1-R^2)/(n-k)}=\mathrm{Eq.}\ (7.50)$$

8.14 (a) $\hat{Y}_t$ =38.969 0+0.2609X_t

se=(3.856 4)　(0.016 7)　r^2=0.942 3

ln $\hat{Y}_t$ =1.404 2+0.589 0lnX_t

se=(0.1568)(0.0293)　r^2=0.964 3

ln $\hat{Y}_t$ =3.931 6+0.002 8X_t

se=(0.046 4) (0.000 2)　r^2=0.924 8

$\hat{Y}_t$ =-192.97+54.213 lnX_t

se=(16.380) (3.062 3)　r^2=0.954 3

(b) 模型(1)中,斜率给出了X每变动一个单位，Y均值的绝对变动。模型(2)中，斜率给出了Y对X的弹性。模型(3)中，斜率给出了X每变动百分之一，Y的增长率。模型(4)中，斜率给出了X的相对变动引起的Y的绝对变动。

(c) 0.260 9, 0.589 0(Y/X), 0.002 8(Y), 54.21 3 (1/X)。

(d) 0.2609(X/Y), 0.5890, 0.0028(X), 54.213(1/Y)。

(e) 上述模型都给出了相近的答案，所以你可以自由选择。

9.11 (a) 销售额$_t$ =4 767.8+912.25D_{2t}+1 398.8D_{3t}+2 909.8D_{4t}

t=(14.714)　(1.990 7)　(3.052 3)　(6.349 6)　R^2=0.779 0

(b) 第一季度平均销售量为4 768百万美元。第二季度、第三季度、第四季度的平均销售量分别比第一季度高912、1399、2910百万美元。各个季度的截距值分别为4 767.8，5 680.05，6 166.6和7 677.6。

10.13 (a) 不是，因为X_3=2X_2-1，也即完全共线性。

(b) 估计：Y_i=A_1+A_2X_{2i}+μ_i，其中，A_1=(B_1-B_2)，A_2=(B_2+2B_3)。回归结果为：$\hat{Y}_t$ =-12.0+2X_{2i}。因此，(B_1-B_2)=-12, (B_2+B_3)=2。

10.20 (a) 第一个模型的斜率系数就是弹性。第二个模型 log$K$和log$H$的系数为弹性。模型中趋势变量的系数表示Q的年增长率为2.7%。

(b) 在零假设：总体系数为零下，估计的t值分别为-3.6、10.195 4和6.518。在5%的显著水平下(双边检验)是显著的，因为临界的t值为2.101。

(c) 趋势变量和log$K$的t值分别为1.333 3和1.381 4。

(d) 可能由于趋势变量与资本共线性。

(e) 即使是高度相关的，也不一定表明是共线性的。

(f) 利用R^2值和F检验得到F=45.384 4，其中分子和分母自由度分别为3和17。在1%的显著水平下，临界的F值为5.18。计算的F值超过了临界值，所以拒绝零假设。

(g) 规模报酬为(0.887+0.893)=1.78，也即规模报酬增加。

11.6 在模型Y_i=B_1+B_2X_i+μ_i中，等式两边同除以X_i^2得到：

(Y_i/X_i^2)=B_1(1/X_i^2)+B_2(1/X_i)+v_i。在这个模型中，v_i是同方差的。

11.11 (a) 令Y=GDP的增长率(%)，X=投资/GDP(%)。Y对X回归。

(b) 是。因为样本中不同的个国家经济背景不一样。

11.14 (a) $\hat{Y}_t$ =1 992.3+0.233 0X_i

t=(2.127 5)　(2.333 3)　r^2=0.437 5

(b) ($\hat{Y}_t$ /σ_i)=2 416.9 (1/σ_i)+0.180 1(X_i/σ_i)

t=(2.1109)　(1.4264)　r^2=0.648 1

注：两个r^2值不能比较。

12.11 都存在自相关。

12.14 **第一阶段回归结果：**

$$\hat{Y}_t = -89.967 + 0.474\,1X_t - 0.412\,7X_{t-1} + 0.795\,2Y_{t-1}$$

t=(-1.951 3) (5.599 1) (-4.322 5) (4.578 4) R^2=0.976 9

估计的ρ值为0.795 2。

第二阶段回归结果：

$$\hat{Y}_t = -320.63 + 0.277\,4X_t$$

t=(-4.971 4) (9.279 8) r^2=0.968 7

注：这是变形后的Y**和**X。

12.21 **分子分母同除以**n^2**，得到**

$$\hat{\rho} = [(1 - d/2) + k^2/n^2]/(1 - k^2/n^2)$$

当n**趋于无穷大时，** $\hat{\rho} = (1 - d/2)$。

13.12 (a) $\ln \hat{Y}_t = -8.401\,0 + 0.673\,1 \ln X_{2t} + 1.181\,6 \ln X_{3t}$

t=(-3.091 2) (4.395 2) (3.912 1) R^2=0.982 4

(b) $\ln \hat{Y}_t = 2.164\,1 + 1.241\,2 \ln X_{2t}$

t=(4.9023) (17.678) R^2=0.9601

因为遗漏了资本变量，所以估计得到的产出-劳动弹性是有偏的，利用式(13-3)**可以证明这个弹性向上偏**1.259 3。

14.12 $PCE_t = -215.12 + 1.007\,0PDI_t$

t=(-6.259) (63.725) r^2=0.996 1

$PCE_t = -240.00 + 1.038\,5PDI_t - 0.021\,4PCE_{t-1}$

t=(-6.668 3) (42.765) (-1.646 8) R^2=0.996 6

(b) **短期**MPC**为**1.038 5。**长期的**MPC**为**1.016 7。**因为滞后的**PCE**是统计不显著的，所以没有分布滞后影响，也就是说，长期和短期的** MPC**几乎相同。**

14.15 **可以用**logit**模型。结果如下：**

$\ln(p_i/1 - p_i) = -4.833\,8 + 3.060\,9 \ln X_i$

t=(-10.860) (11.875) r^2=0.979 2

回归结果表明：剂量每增加1**个百分点，平均而言，死亡的可能性增加** 3.06**个百分点。**

15.16 (a) **简化形式的回归模型为：**

$$Y_{1t} = \prod_{10} + \prod_{11} X_{1t} + \prod_{12} X_{2t} + v_{1t}$$

$$Y_{2t} = \prod_{20} + \prod_{21} X_{1t} + \prod_{22} X_{2t} + v_{2t}$$

(b) **利用阶条件，两个方程都是恰度识别的。**

(c) **用间接最小二乘法。因为它给出了结构系数的一致估计值。**

(d) **第一个方程是可识别的，第二个是不可识别的。**

15.17 (a) **系数**B_1,B_2**可由简化形式的系数估计。估计值分别为-** 5**和**1.5。

(b) (1) **若**A_2**为**0**，则没有联立问题。因此，两个方程都可以估计。**

(2) **若**A_1**为**0**，两个方程都可以估计。**

参考文献

初级(Introductory)

Frank, C. R., Jr.: *Statistics and Econometrics*, Holt, Rinehart and Winston, Inc., New York, 1971.

Hu, Teh-Wei: *Econometrics: An Introductory Analysis*, University Park Press, Baltimore, 1973.

Katz, David A.: *Econometric Theory and Applications*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1982.

Klein, Lawrence R.: An *Introduction to Econometrics*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

Walters, A. A.: *An Introduction to Econometrics*, Macmillan & Co., Ltd., London, 1968.

中级(Intermediate)

Aigner, D. J.: *Basic Econometrics*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

Dhrymes, Phoebus J.: *Introductory Econometrics*, Springer-Verlag, New York, 1978.

Dielman, Terry E.: *Applied Regression Analysis for Business and Economics*, PWS-Kent Publishing Company, Boston, 1991.

Draper, N. R. and H. Smith: *Applied Regression Analysis*, 2d ed., John Wiley & Sons, Inc., New York, 1981.

Dutta, M.: *Econometric Methods*, South-Western Publishing Company, Incorporated, Cincinnati, 1975.

Goldberger, A. S.: *Topics in Regression Analysis*, The Macmillan Company, New York, 1968.

Gujarati, Damodar N.: *Basic Econometrics*, 3d ed., McGraw-Hill, New York, 1995.

Huang, D. S.: *Regression and Econometric Methods*, John Wiley & Sons, Inc., New York, 1970.

Judge, George G., Carter R. Hill, William E. Griffiths, Helmut Liltkepohl, and TsoungChao Lee: *Introduction to the Theory and Practice of Econometrics*, John Wiley & Sons, Inc., 1982.

Kelejian, H. A. and W. E. Oates: *Introduction to Econometrics: Principles and Applications*,

2d ed., Harper & Row, Publishers, Incorporated, New York, 1981.

Koutsoyiannis, A.: *Theory of Econometrics*, Harper & Row, Publishers, Incorporated, New York, 1973.

Mark, Stewart B. and Kenneth F. Wallis: Introductory *Econometrics*, 2d ed., John Wiley & Sons, Inc., New York, 1981. A Halsted Press Book.

Murphy, James L.: *Introductory Econometrics*, Richard D. Irwin, Inc., Homewood, Ill., 1973.

Netter, J. and W. Wasserman: *Applied Linear Statistical* Models, Richard D. Irwin, Inc., Homewood, Ill., 1974.

Pindyck, R. S. and D. L. Rubinfeld: *Econometric Models and Econometric Forecasts*, 4th ed., McGraw-Hill Book Company, New York, 1998.

Sprent, Peter: *Models in Regression and Related Topics*, Methuen & Co., Ltd., London, 1969.

Tintner, Gerhard: *Econometrics, John Wiley & Sons*, Inc. (science ed.), New York, 1965.

Valavanis, Stefan: *Econometrics: An Introduction to Maxim u m-L ikelihood Methods*, McGraw-Hill Book Company, New York, 1959.

Wonnacott, R. J. and T. H. Wonnacott: Econometrics, 2d ed., John Wiley & Sons, Inc., New York, 1979.

高级(Advanced)

Chow, Gregory C.: *Econometric Methods*, McGraw-Hill Book Company, New York, 1983.
Christ, C. F.: *Econometric Models and Methods*, John Wiley & Sons, Inc., New York, 1966.

Dhrymes, P. J.: *Econometrics: Statistical Foundations and Applications*, Harper & Row, Publishers, Incorporated, New York, 1970.

Fomby, Thomas B., Carter R. Hill, and Stanley R. Johnson: *Advanced Econometric Methods*, Springer-Verlag, New York, 1984.

Gallant, Ronald A.: *An Introduction to Econometric Theory*, Princeton University Press, Princeton, New Jersey, 1997.

Goldberger, A. S.: *Econometric Theory*, John Wiley & Sons, Inc., New York, 1964.

Goldberger, A. S.: A *Course in Econometrics*, Harvard University Press, Cambridge, Mass., 1991.

Greene, William H.: *Econometric Analysis*, Macmillan Publishing Company, New York, 1990.

Harvey, A. C.: *The Econometric Analysis of Time Series*, 2d ed., MIT, Cambridge, Mass., 1990.

Johnston, J.: *Econometric Methods*, 3d ed., McGraw-Hill Book Company, New York, 1984.

Judge, George G., Carter R. Hill, William E. Griffiths, Helmut Liltkepohl, and TsoungChao Lee: Theory and Practice of Econometrics, John Wiley & Sons, Inc., New York, 1980.

Klein, Lawrence R.: *A Textbook of Econometrics*, 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1974.

Kmenta, Jan: *Elements of Econometrics*, 2d ed., The Macmillan Company, New York, 1986.

Madansky, A.: *Foundations of Econometrics*, North-Holland Publishing Company, Amsterdam, 1976.

Maddala., G. S.: *Econometrics*, McGraw-Hill Book Company, New York, 1977.

Malinvaud, *E.: Statistical Methods of Econometrics*, 2d ed., North-Holland Publishing Company, Amsterdam, 1976.

Theil, Henry: *Principles of Econometrics*, John Wiley & Sons, Inc., New York, 1971.

专业(Specialized)

Belsley, David A., Edwin Kuh, and Roy E. Welsh: *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons, Inc., New York, 1980.

Dhrymes, P. J.: Distributed Lags: *Problems of Estimation and Formulation*, Holden-Day, Inc., Publisher, San Francisco, 1971.

Goldfeld, S. M. and R. E. Quandt: *Nonlinear Methods of Econometrics*, North-Holland. Publishing Company, Amsterdam, 1972.

Graybill, F. A.: *An Introduction to Linear Statistical Models*, vol. 1, McGraw-Hill Book Company, New York, 1961.

Rao, C. R.: *Linear Statistical Inference and Its Applications*, 2d ed., John Wiley & Sons, New York, 1975.

Zellner, A.: *An Introduction to Baycsian Inference in Econometrics*, *John* Wiley & Sons, Inc., New York, 1971.

应用(Applied)

Berndt, Ernst R.: *The Practice of Econometrics: Classic and Contemporary*, Addison-Wesley, 1991.

Bridge, J. I.: *Applied Econometrics*, North-Holland Publishing Company, Amsterdam, 1971.

Cramer, J. S.: *Empirical Econometrics*, North-Holland Publishing Company, Amsterdam, 1969.

Desai, Meghnad: *Applied Econometrics*, McGraw-Hill Book Company, New York, 1976.

Kennedy, Peter: *A Guide to Econometrics*, 3d ed., MIT Press, Cambridge, Mass., 1992.

Leser, C. E. V: *Econometric Techniques and Problems*, 2d ed., Hafner Publishing Company, Inc., 1974.

Rao, Potluri and Roger LeRoy Miller: *Applied Econometrics*, Wadsworth Publishing Company, Inc., Belmont, Calif., 1971.

Note: For a list of the seminal articles on the various topics discussed in this book,please refer to the extensive bibliography given at the end of the chapters in Fomby et al., cited previously.