目　　录

第一章　引　论

1.1　计量经济学的性质

计量经济学是一门包含经济学、数学、统计学三者的方法论方面的学科。计量经济学者就是试图运用一些主要根据统计学方法论制定的、并且通常用数学语言表达的技巧来理解经济系统运行的经济学者。这一学科庄严的背景可能有时会使那些愿意了解计量经济学道理的人望而却步。本书的目的，就是试图解释为什么要使用计量经济学的方法以及如何使用计量经济学的方法，尽可能避免使用高深的数学形式，但对有关的统计概念则要加以详细的解释。虽然引入的概念是严密的并且必须小心运用，但按部就班的数学推导和证明并不一定是必需的，而解释某个特定结论何以能够成立往往更为必要。

1.2　经济模型

在开展进一步研究之前，必须准确地了解一个**经济模型**指的是什么。经济系统无疑是复杂的，正是由于这种复杂性，产生了使用模型的想法。模型是对现实的抽象，它要展现出该经济系统的主要特点。显然，模型可以是“好”的或是“坏”的。假如抽象过分，这个模型就说明不了它想说明的现实系统，相反，如果抽象不足，这个模型就会太复杂，对洞察现实系统的运行就没有什么价值。

模型可以有许多种形式。对任何系统的分析都必须以一个模

型为基础，但这个模型不一定是明显的。经济报刊提供许多分析的实例，它们显然都是建立在一组假设条件之上的，这些假设条件有些是明显的，但常常不是这样，它们代表着一个隐含的模型。假如模型能以较为明显的形式给出，对那些想对报刊的分析作出评价的人显然更为有利。

计量经济学者所用的模型一般是以数学形式表达出来的。但一般的明显的模型，或某些特殊的经济模型，并不用数学形式表达。一个表示全国经济中货物、劳务和资金流量的图形是一个模型，但一个模型一定要适于经济学者所要了解的问题，如果他关心的是这些流量之间的**关系**，流量图本身就不能满足需要。因为在这种情况下，抽象的程度太过分了。

本章其余部分的讨论主要是以公认的宏观经济中消费者支出与个人可支配收入的关系为例。主张这样一种关系应该存在的断言本身就是一个模型，但它是一个不够精密的模型，不足以说明消费的变化对收入的变化所作反应的程度。要进一步说明问题，就必须给这种关系以比较明显的形式。办法之一就是从消费者个人对其收入的变化所作反应的**理论**出发，在把许多个人行为的**总和**考虑在内之后，就能得到宏观经济中消费与收入之间的关系的理论说明。这个理论的发展可以提供表达这一关系的确切形式，但这常常不能做到，因此经常采用的原则是使关系式尽可能简化，直至根据对现实系统的观察，有证据表明那个简单的形式已经到了站不住脚的时候为止。由此引起一个有趣的观点：一个模型是对现实的抽象，而一个“正确的”模型并不需要对现实系统的运行作出完整无缺的描述。只要现实系统中决策者的行为**好象**是在遵循着模型的规则，那就够了。但是，假如模型所含的某些特定的假设条件在实践中受到干扰，这个模型就不会正确；假如任意一个假设条件**看来**已受到干扰，这个模型**看来**也就不会正确。对假设条件的检验是应用计量经济学技术的一个重要问题，

在以后适当的章节中，将更多地谈到这个问题。

1.3 简单模型

设若一个假想的经济中消费者支出（C）和个人可支配收入（D）之间存在一种关系，它可以表达为：

$$C = 50 + 0.8D \qquad (1.3.1)$$

在这个方程中，C和D是**变量**，它们在经济系统的不同观测点可以取不同的值。数字50和0.8是常数。为了集中注意力，假定我们是对两个时期中每一时期的消费流量和收入流量进行观察。那么，虽然每个时期的消费和收入一般会取得明显不同的值，但是一种固定关系的存在，却意味着数字50和0.8仍然是一样的。虽然变量的值会有变化，但变量之间的关系并没有变。

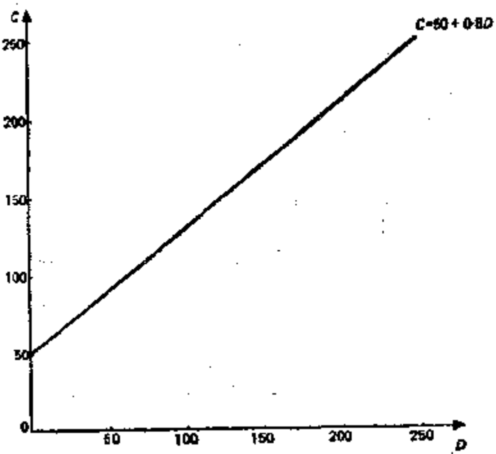

图1.1

图1.1是表示方程（1.3.1）的图形。可以看到，该方程

在图中表现为一条直线。如果在这假想的经济中消费行为可用这个方程来描述，那么，每一对消费和收入的值都会表现为直线上的一点。从图中也可以看到，如果收入为零，消费的值按方程就是50。如果收入增加5个单位，消费就会上升4个单位。同样，如果收入增加1个单位，消费就会增加4/5=0.8单位。因此数字0.8表现了**释变量**（explanatory variable）（收入）每发生1个单位的变动对应变量（dependent variable）（消费）所产生的影响。方程（1.3.1）是直线或**线性**关系的一个特例。更一般地，C和D之间的线性关系可以写成

$$C=\alpha+\beta D \tag{1.3.2}$$

这里α和β是这个关系式的**参数**。在我们假想的经济中，$\alpha=50$，$\beta=0.8$。在另一种情况下，线性关系也可以再次形成，但参数值却可能不一样。参数α称为**截距**，β称为**斜率**。这里对参数的解释和以上对数字50和0.8的解释是严格一致的。为了从代数的观点看出β表示的是斜率，假定两个时期变量的值可写成C_1，D_1和C_2，D_2。于是，两个时期之间D的**变动**为D_2-D_1，而C的变动为C_2-C_1。假如消费行为象（1.3.2）式所描述的那样，每一对变量都能满足方程，于是有

$$C_1=\alpha+\beta D_1 \tag{1.3.3}$$

$$C_2=\alpha+\beta D_2 \tag{1.3.4}$$

如果这两个方程都正确，那么，这两个方程相加或相减所得的方程也正确。从（1.3.4）式减去（1.3.3）式就得到把消费**变动**和收入**变动**联系起来的方程

$$(C_2-C_1)=0+\beta(D_2-D_1)$$

因此有

$$\beta=(C_2-C_1)/(D_2-D_1) \tag{1.3.5}$$

如果现在我们考虑收入发生1个单位的变动，即$(D_2-D_1)=1$，

参数β就表示消费所发生的相应的变动。应该注意，一个单位收入的变动无论从哪一个特殊的点发生，对收入一个单位的变动，消费的变动总是等于β。

在经济术语中，上述关系式是**消费函数**的线性表述。β是**边际消费倾向**。边际消费倾向就是消费函数的斜率，但必须认识，只有在线性函数的情况下，斜率才是一个不依赖于C或D的取值的常数。与**非线性**关系式对应的图形是曲线，而不是直线。在曲线的情况下，斜率就会随变量的值变动而变动。

下面我们将主要集中注意于线性关系，而在一些不同的段落中，认清给出的关系何时是线性关系，这是十分重要的。我们可以绘制与这种关系相应的图形，看它象不象一条直线。但一种更简单的检验方法是看这种关系能否写成方程（1.3.2）的形式。例如

$$5C = 250 + 4D$$

所表现的就是线性关系。因为两边除以5，即得

$$C = 50 + 0.8D$$

这是标准的线性形式。对比之下，方程

$$C = 20 + 10\sqrt{D}$$

则表示变量C、D之间的关系是**非线性的**。因为无法把它写成（1.3.2）那样的方程。即使会使讨论复杂化，我们还是要进一步说，在后一个例子中，变量C和$\sqrt{D}$是一种线性关系。因为如果我们画图时在横轴上所标明的是$\sqrt{D}$的值而不是D的值，这个方程就会画成一条直线。但如果对D而言，C的图形则是一条曲线。这可以用图1.2来说明。

假如在消费与收入之间存在确切的线性关系，那么，通过对两个时期流量的观察，在图上描出两对数值并联结两点，这个关系就可获知。这就是说，α和β的数值可以获知。用这图形或用写有α和β的适当数值的公式，就可以计算出在任意给定的收入

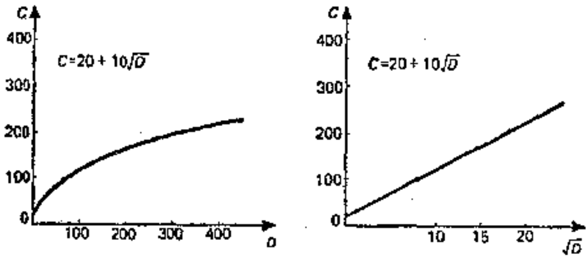

图1.2

水平下的消费水平。但对任何现实的经济，如果我们取许多年的消费与可支配收入的数字，并在图上描出许多点，一般是不可能找出一条穿过所有各点的直线的。那么，是不是要采用一个代表曲线的方程来代替直线方程呢？实际上，要画出通过若干点的某种曲线总是可能的，但是，如果再多取一个观测点，我们不可避免地就会发现它并非正好落在我们所选择的曲线上。虽然线性形式并非总是适当的，但选择更复杂的曲线也不是解决问题的办法。

图1.3表示的是根据1963—1972年联合王国经济资料中，按1963年价格计算的消费者支出和与之对应的按同一价格计算的个人可支配收入资料所描的点。由于显然不可能画出通过所有这些点的一条直线，所以要建立一个表明那种可能性存在的模型也是不合理的。现实经济系统中消费与收入之间存在着联系，这是合乎实际的，但表达这种联系的模型一定要改变。因此，我们可以假定，一个不确切的线性关系能够提供一个满意的模型。这个模型可以写成

$$C=\alpha+\beta D+u \qquad (1.3.6)$$

这里，u表示对C、D间关系的一个**扰动因子**。

模型的主要部分如果按现在这样设定，扰动因子所代表的就

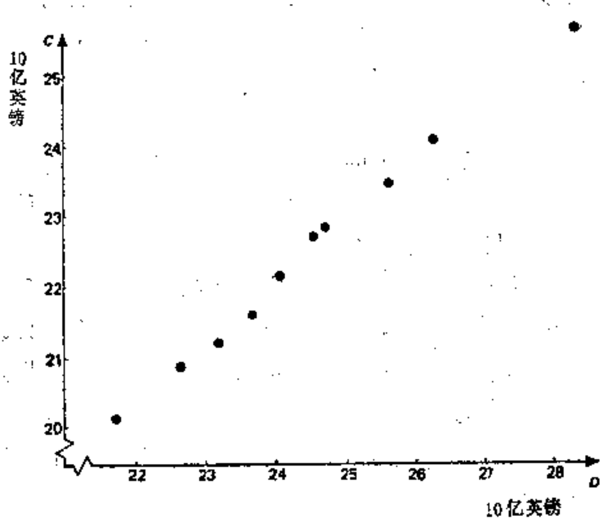

图1.3

是一切可能的误差，而不论这些误差能否避免。这些误差，是使用方程$C=\alpha+\beta D$作为基本经济假设时必然出现的。假如消费还受其他方面的重要影响，干脆把它们当作扰动因子的一部分是可以的。但多用几个释变量可能会更加令人满意，因为这样就可以把重要的影响都明显地包括在模型的主要部分之中。增加释变量个数的方法将在第三章介绍。但是，还可能存在其他一些因素，它们并非特别重要，又不容易测量。因为不可能有其他的表现形式，使用扰动因子就是把这些因素包括进来的一个方便而合理的技巧。这就是说，如果收入是对消费的唯一主要影响，方程（1.3.6）就是可以接受的。假如有其他的重要影响没有包括在模型的主要部分之中，这个方程就不可接受。

记住这点，我们就要对扰动项加上一组条件。这些条件是这样挑选出来的：除非模型的主要部分包括了对消费的所有主要影响，否则这些条件就不能得到满足。模型的主要部分在其他方面

也应该是可以接受的。我们也要说，不用线性方程而用其他形式也需要扰动项。但是，除非线性关系是对实际经济系统中收入与消费间的联系的合理、良好的逼近，否则，加在扰动因子上的条件是不会得到满足的。

要确定应施加于扰动项的条件，就要把建立模型的想法运用到扰动因子数值的确定过程中。这时所用的模型很象那种用于机遇游戏的模型。机遇游戏这个名称，来源于这种游戏事先不可能确定取得什么样的结果。机遇游戏实际上指的是这样一种模型：由于很难了解一个特殊结果**为什么**会产生，就把这个结果说成是由机遇确定的。这给人以一个含混不清的印象，好象结果是以极其随意的方式选出的。但是，为了达到我们的目的，必须建立一个更明显的模型的形式，模拟机遇游戏本身并非我们的兴趣所在，但是它有助于引进某些概念，可以用它们来模拟消费函数中的扰动项。

1.4 机遇游戏模型

有一颗标有从 1 至 6 的数字的六面骰子。掷一次，这个游戏可能的结果是 1， 2， 3， 4， 5， 6 这几个不同的数值。这种游戏所得结果的基本特点是：抛掷一次，只能有一个结果。因此这些结果可以说是**互相排斥的**。还可以再想想某些结果的集合，我们称它构成特定的**事件**。例如，有一种集合是"掷得的点大于 3"。如果掷的结果为 4， 5， 6 三者之一，则这一事件就会发生。还有可能定义**能够**同时发生的两个事件。比如，如果把一个事件定义为"掷得的点大于 3"，第二个事件定义为"掷得的点小于 5"，那末，当骰子掷出"4"时，这两个事件就都发生。可见结果总是互相排斥的，而事件可能是、也可能不是互相排斥的。

当一颗骰子被抛掷，无论情况有多么复杂，都有理由推测会

取得某一特定的结果。但是由于实际的机制十分复杂，用模型的概念来考察游戏容易得多。按照模型，每一结果都与代表该特定结果的机会或机遇的测度相联系。更一般地说，这样一种测度也与由各个结果构成的任一事件相联系。这种测度就是一个事件的**概率**，它有下列性质：

1.一个事件的概率大于（或等于）0。

2.一个必然发生的事件的概率等于1。

3.对于各个互相排斥的事件，其中一个事件或另外各个事件发生的概率是它们各自的概率之和。

从概率的这些**公理**出发，测度的所有其他性质都可以推导出来。但是，对概率这个概念的解释，我们实际需要知道的只是：概率是一个数，它的值位于0与1之间；概率越高，事件发生的可能性越大。

掷骰子一次的模型，它的完整描述是一张列出各种可能的结果以及与之相联系的概率的表。一个特别简单的模型可以根据“所有结果的出现都有同等可能性”的假设作出。在这种情况下，概率测度的性质，会指出每种结果都有1／6的概率。掷骰子后**必定**出现一个结果，因此，掷得1或2或3或4或5或6的概率必定是1。但这些结果是互相排斥的，所以，掷得1或2或3或4或5或6的概率等于这6种结果的概率之和。这样，这6种个别概率之和必定等于1；如果还假定每个结果都有**相等**的概率，那末这个共同的概率一定是1／6。

列出各种可能结果及其相应概率的表称为**概率分布**。可以把各项结果表示为变量V所能取的不同的值。有多少个不同的结果，这个变量就可以取多少个不同的值。因此，对掷一颗骰子来说，变量的值是：$V=1$，$V=2$，$V=3$，$V=4$，$V=5$，$V=6$。对上述每一个值都有一个相应的概率，假设为1／6。一个变量，若其值按某一概率分布的规律而确定，则此变量称为**随机**

变量。所以掷一颗骰子的模型是基于这样的假设，即：一个代表多种可能结果的变量可以看作一个随机变量。

上面所举的例子是很明确的：可能结果的数目很少，假定它们各自的概率相等也很有道理，因此，在这个基础上很容易确定一个完整的概率分布。我们将运用类似的原则对消费函数的扰动因子建立一个模型。但是，随机变量的某些性质，最好通过比较简单的例子加以介绍，为此，下面我们将继续讨论抛掷骰子的游戏。

第一步考虑随机变量的**期望**或**期望值**。为了求得期望值，就要把随机变量的每一个可能值乘以相应的概率，然后将所得乘积相加。对一个随机变量V，其期望值写为$E(V)$，这样，对于掷一颗骰子，有

$$E(V)=1/6(1)+1/6(2)+1/6(3)+1/6(4)+1/6(5)+1/6(6)=3.5$$

实际上，3.5这个数值永远不会作为游戏中的一个结果而出现。单独一次抛掷只会出现从$V=1$到$V=6$的六种结果之一。期望值所表达的只是所有可能出现的数值的一个平均数。一般说来，它并不是将所有结果的值简单相加后除以结果的个数而得。因为有些结果可以多于或少于其他的结果。为了取代这种办法，我们在计算中对每个结果都用相应的概率加权。在这个掷骰子的例子中，每一个结果都有同样的概率，各个结果的简单平均数就可以得出正确的期望值。

一个随机变量的期望值有时称为相应的概率分布的**均值**，或**总体均值**。这些名称表达的完全是同一个概念，它们可以互换。重要的问题是要认清，$E(V)$本身不是一个随机变量，它是传递关于随机变量V的一部分信息的一个常数值。

现在谈谈**方差**的概念。对一个随机变量V，其方差写为

var (V)，并定义为

$$\text{var}(V)=E([V-E(V)]^2) \qquad (1.4.1)$$

这个表达式的解释并不是十分明白易晓的，因此在试图解释方差代表什么之前，我们先来计算掷骰子游戏的方差。我们知道$E(V)=3.5$，所以$[V-E(V)]$能取下列各值：

$1-3.5=-2.5$；　$2-3.5=-1.5$；　$3-3.5=-0.5$；

$4-3.5=0.5$；　$5-3.5=1.5$；　$6-3.5=2.5$。

现在考虑$[V-E(V)]^2$的可能值，

如果$V=1$或$V=6$，则$[V-E(V)]^2=6.25$

如果$V=2$或$V=5$，则$[V-E(V)]^2=2.25$

如果$V=3$或$V=4$，则$[V-E(V)]^2=0.25$

因为$V=1$和$V=6$这两个结果是互相排斥的，所以获得“$V=1$或$V=6$”的概率是其各自的概率之和，于是得1/3。因此，$[V-E(V)]^2=6.25$的概率也是1/3。同理，$[V-E(V)]^2$其他两值的概率也为1/3。

刚才我们说明的是$[V-E(V)]^2$所能取的某些值以及取这些值的一定的概率，换言之，$[V-E(V)]^2$是一个随机变量。由于我们知道这个随机变量的适当的概率分布，因此我们就能算出它的期望值：

$$E[V-E(V)]^2=\frac{1}{3}(6.25)+\frac{1}{3}(2.25)+\frac{1}{3}(0.25)=2.92$$

按照(1.4.1)的定义，这个数值就是V的方差。

现在对方差var (V)作些解释。$E(V)$表达的是V可能出现各值的平均数，方差var (V)则将各个数值与其期望值的差异的程度综合为一个单独的测度。因此，方差是一种离散的测度(a measure of spread)。对每一个结果来说，数量$[V-E(V)]$

表示的是它与期望值的离差。为了求得离差的平均值，人们可能会想到要取离差的期望值。但有些离差是正数，有些离差是负数，离差的期望值实际上为零（见习题1.2）。相比之下，所有的离差的**平方**都是正数，离差平方的期望值也是正数。采用将离差平方的办法，我们就可以保证正、负离差对离散测度的形成都能起同样的作用。

我们已经看到，在掷骰子的游戏中，方差var（V）＝2.92，现在知道这是一个离散的测度，而2.92这个数值实际上告诉我们什么仍然不完全清楚。在习题1.1中，要求读者考虑游戏的两种变化形式。第一种变化形式是，骰子面上分别标有数字7，8，9，10，11和12。在这种情况下，期望值增大了，但是方差并没有变，因为数值的离散度并不比原来的游戏大。在第二种变化形式中，骰子上所标的数字是2，4，6，8，10和12。在这种情况下，期望值和方差都增大了。这说明了一个重要的问题：我们可以把从两个不同的随机变量得来的方差加以比较，也可以将一个随机变量的方差与其他数值进行比较，但是，我们通常并不孤立地考虑方差。因此，重要的并不是2.92这个数字本身，而是这个数字和其他一些数值的比较，这种比较将告诉我们一些重要的东西。

现在假定游戏由骰子的两次抛掷组成。掷骰子两次的结果可用两个随机变量V_1和V_2来表示。V_1的取值表示掷第一次的结果，V_2的取值表示掷第二次的结果。这意味着现在有两个概率分布，但由于没有理由认为在两次抛掷之间，实际系统的机制会有什么不同，我们可以假定V_1和V_2有同样的一组结果和同样的一组概率。这样，这两个分布就可以说是完全相同的。这句话的含义之一是，这两个随机变量具有相同的期望值和相同的方差。也没有理由认为在实际系统中，第一次抛掷所得的值会对第二次抛掷的值发生什么影响。这在模型中的反映，就是“这两个分布是**独立**

的”这样的假设。对于独立这个概念我们无需给出精确的专门的定义，但要注意它的一个重要的含义。在分析两个独立的随机变量的行为时，我们能为每一个变量设定它的分布，这就提供了一个完整的描述。没有必要考虑两个变量的相互影响，因为独立意味着没有相互影响存在。

这个道理马上就可以推广到一个由抛掷骰子n次构成的游戏上。其模型会包括一组n个随机变量，V_1，V_2，…，V_n，每一个随机变量都有相同的分布，而每一个分布都是独立于所有其他分布的。这种推广有一个重要的理由。到目前为止，我们所谈的都只是关于掷一颗骰子的理论模型，但是，很显然，通过实际掷一颗骰子，并把掷的结果记录下来，我们也可以从相应的实际系统取得观测值。把这些资料搜集起来，我们可以得到一组来自实际系统的观测值和一个可能用以“解释”观测值的模型。从原则上说，在这组观测值的基础上，我们就可以在实际系统和模型之间进行比较。单独掷一次的模型没有多大用处，因为实际系统中的相应观测值也只能是掷一次的结果，它能告诉我们多少关于实际系统的情况是很可疑的。比较常见的情况是取一组观测值，而要进行比较，更需要的是能表达一组观测值生成过程的模型。

现在我们有一个模型，它在许多方面都和用于消费函数扰动项的模型相似。然而，还有一个小问题需要解释。我们将会从下一节的讨论中看到，从实际经济系统中取得的资料是由消费和收入的一组观测值构成的。这些观测值通常是与不同的时期相联系的。消费的n个观测值的自然表达方式是$C_1, C_2, \cdots$，C_n。但是，根据模型，我们也将发觉消费的每一个观测值都正好是相应的随机变量可能取值中的一个。这里，决定性的区别在于**变量**和**随机变量**这两个术语的用法。一种可能的混淆来自运用概率模型的学科对变量一词的用法。消费是一个单独的经济变量，但我们可以有一个模型，这个模型就意味着消费的每一个观测值都对应于一

个不同的随机变量。为了避免形成一套完全不切实际的符号系统，正如V_1，V_2，…，V_n在抛掷骰子的模型中是一组随机变量那样，我们也用C_1，C_2…，C_n等符号来表示一组随机变量。能做到这一点就不会引起混乱，而联系上下文看，我们想说的是什么也就很清楚了。如果在一个对观测数据进行计算给以指令的公式中出现符号C_1，C_2…，C_n，在实际运算时，它们就会被实际观测值所代换。但在讨论模型的理论含义时，我们亦可取C_1，C_2…，C_n来表示一组随机变量。

1.5 随机扰动因子

现在把前一节所谈的概念应用于消费函数的扰动项。在方程（1.3.6）中提出的模型是：

$$C = \alpha + \beta D + u \qquad (1.5.1)$$

但在实践中，这模型会被应用于消费和收入的一组特定的观测值。这些值通常涉及不同时期的某国经济，这就要假设某一段时间内消费函数保持不变。我们也可以说，某些不同的经济具有同样的消费函数，从而对不同的地理区域取观测值。无论是哪一种情况，消费的一组n个观测值都可以写成C_1，C_2…，C_n，或更简捷地写成C_t（$t=1$，2，…，n）。相应的收入观测值可以写成$D_t(t=1,2,\cdots,n)$。这样，用于这些观测值的模型就是：

$$C_t = \alpha + \beta D_t + u_t; \quad t=1,\ 2,\ \cdots,\ n \qquad (1.5.2)$$

它指出，在一定的常数α和β的情况下，对n个时期如$t=1$，$t=2$，等等，该模型成立。我们通常谈的是时期，但应了解这一论证同样适用于任何其他类型的观测值。与一个单独的单位相联系的不同时期的一组观测值称为**时间序列**。在某一个时点上的不同单位的一组观测值是一个**横截面**。在不同的场合，这些单位可以是户，公司，行业，地区或国家。但至少是在宏观经济中，时间序列是最常用的观测值，我们的讨论将继续以此为基础。

现在假定每个时期扰动因子的行为好象是遵循着概率分布规律的。在最简单的情况下，也可假定所有的分布都是相同的，而且它们都是独立的。因此，在一组n个扰动因子的模型和骰子n次抛掷的模型之间有一个很明显的类似之处。骰子的每一次抛掷都有同样的概率，而且任何一次抛掷的结果对其他次抛掷的结果都没有影响。类似地，同样的概率适用于每个时期扰动值的生成，而在一个时期取得的扰动值对其他时期取得的扰动值并无影响。

迄今我们仍未谈及与每一个扰动项相联系的概率分布的性质。事实上，我们只需要考虑一个单独的扰动因子，比如说u_1，因为已经假定所有其他的扰动因子都有相同的分布。图1.4说明我们将要应用的分布。

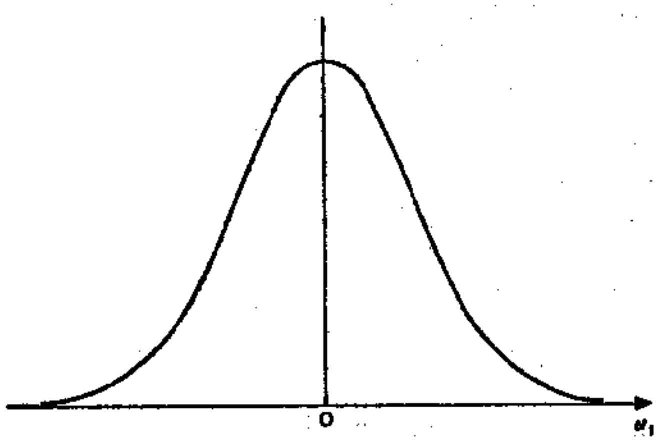

图1.4

横轴是一条表示u_1各可能值的直线，为了满足这个分布的要求，一定要把u_1看成一个连续随机变量，它能取一个与这条线上无数点中任何一点相对应的值。在横轴之上是一条曲线，它