华章教育

PEARSON Education

经济教材译丛

中文版

现代计量经济学

下册

Econometrics: A Modern Introduction

（美） 迈克尔 P. 莫瑞 (Michael P. Murray) 著
贝茨学院

费剑平 译

机械工业出版社
China Machine Press

经济教材译丛

中文版

现代计量经济学

下册

Econometrics: A Modern Introduction

(美) 迈克尔 P. 莫瑞 (Michael P. Murray) 著
贝茨学院

费剑平 译

机械工业出版社
China Machine Press

本书系统介绍了计量经济学理论及其在经济实践中的应用。不仅注重计量经济学理论和方法的经典文献，还注重计量经济学的理论和方法在其他经济学应用研究中所取得的重要成果，包括一些有意思的案例。在讲解计量经济学理论和方法的过程中，试图用语言文字把道理讲得透彻，而不是依靠数学推导。本书是下册内容，构成了计量经济学第二学期课程的基础。

Michael P Murray. Econometrics：A Modern Introduction.

ISBN 0-321-11361-6

本书版权登记号：图字：01-2007-2285

图书在版编目（CIP）数据

现代计量经济学（下册）/（美）莫瑞（Murray，M. P.）著；费剑平译. —北京：机械工业出版社，2009. 3

（经济教材译丛）

书名原文：Econometrics：A Modern Introduction

ISBN 978-7-111-26624-2

Ⅰ. 现…　Ⅱ. ①莫…　②费…　Ⅲ. 现代计量经济学－教材　Ⅳ. F224. 0

中国版本图书馆 CIP 数据核字（2009）第 038493 号

机械工业出版社（北京市西城区百万庄大街 22 号　邮政编码　100037）

责任编辑：胡智辉　　　　版式设计：刘永青

北京京北印刷有限公司印刷

2009 年 3 月第 1 版第 1 次印刷

184mm × 260mm · 18. 25 印张

标准书号：ISBN 978-7-111-26624-2

定价：46. 00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

本社购书热线：（010）68326294

投稿热线：（010）88379007

目录

下　　册

第12章

估计量的大样本性质：一致性和渐近有效性

只要将数学定律用于现实，
它们就是不确定的，
而只要它们是确定的，
那它们就无法用于现实。[1]

——爱因斯坦

一个星期六的夜晚，一名醉汉在路灯下用双手和双膝着地缓慢地爬行着。一名警察碰巧经过。

“你在干什么？”警察盘问道。

“我在找我的车钥匙。”醉汉含糊不清地回答道。

“哦，你把它掉在这里了，是吗？”警察问道。

“不是的，我把它扔在那里了，”醉汉一边指向一个黑暗的小巷一边回答道，“但这里的光线好一些。”

醉汉的办法向我们警示了计量经济建模中的一个风险。简化假定和在分析上对方便的让步可能会破坏经验研究的基础，让你从一开始就注定失败。不过，巧妙选择的简化假定和分析便利性对于促进我们的理解有重要帮助。

无偏性和一致性是在直觉上有吸引力的估计量选择准则，但在以后章节将遇到的数据生成过程中，寻找无偏估计量通常都倍受打击。要么寻找一个无偏估计量在数学上超出了我们的能力，要么我们不具备无偏估计量所需要的信息。为了克服这些困难，计量经济学家在新的领域寻找最佳估计量。他们设计无偏性和有效性的便于分析的替代选择，作为选择估计量的准则。在这些新的分析中，计量经济学家研究估计量随着样本容量无限增大时的表现。

本章首先阐释的一个大样本性质是计量经济学家最感兴趣的一致性，然后我们来看前面章节中研究过的无偏估计量是否也是一致估计量。本章的一个核心结论是，在解释变量在不同样本中并非保持不变的现实情形中，最小二乘估计量有可能是一致的，也有可能是不一致的。就像有很多无偏估计量一样，在一个给定的数据生成过程中，可能有许多估计量都是一致的，本章接下来就介绍渐近有效的概念，我们可以利用渐近有效性在很多备选的一致估计量中进行选择。判断一个估计量的渐近有效性，要求我们判断估计量的分布随着样本容量无限变大时的变化。

胡乱使用一致性和渐近分布的概念，就像醉汉在街灯下寻找车钥匙一样具有误导性，但明智地使用这些概念，在没有无偏估计量可用时，可能会让我们得到一些好的估计量，并因而能够进行稳靠的推断。

12.1 大样本或渐近的视角

到目前为止，我们一直考虑的是，从一个给定容量的许多样本得到的估计值的性质：“一个估计量是无偏的吗？它的方差是多少?”我们想知道哪个估计量在一个容量为 n 的样本中具有“最好的抽样性质”。本章的新视角为我们提供了另一个考虑。它研究估计量的**大样本性质**(large-sample properties)或**渐近性质**(asymptotic properties)，即估计量随着样本容量无限扩大时所具有的性质。

在我们放松了高斯—马尔科夫假定或者把线性于参数的模型变成非线性模型时，无偏性和有效性通常都变成无法实现的目标。在寻找具有良好的跨样本性质的估计量倍受挫折时，计量经济理论家常常转而推导具有良好的大样本性质的估计量，并希望那些在很大的样本中表现优良的估计量在中等大小的样本中也能很好地发挥作用，而且这种愿望常常都能实现。

既然我们的样本通常远远达不到无限，那么我们为什么满足于那些适用于接近无限大样本的性质呢？这是因为，确定一个估计量的无偏性和有效性等抽样性质，有时候太困难了。我们之所以转向大样本性质，是因为大样本性质通常在数学上比有限样本性质更容易分析。

不过，这个新的视角并没有完全放弃原来的视角。研究大样本性质的背景仍是通过对许多样本的考查，只不过，这个新的视角不再考查给定容量下不同样本的期望值，而是考虑随着样本容量无限增大出现有利结果和不利结果的概率。

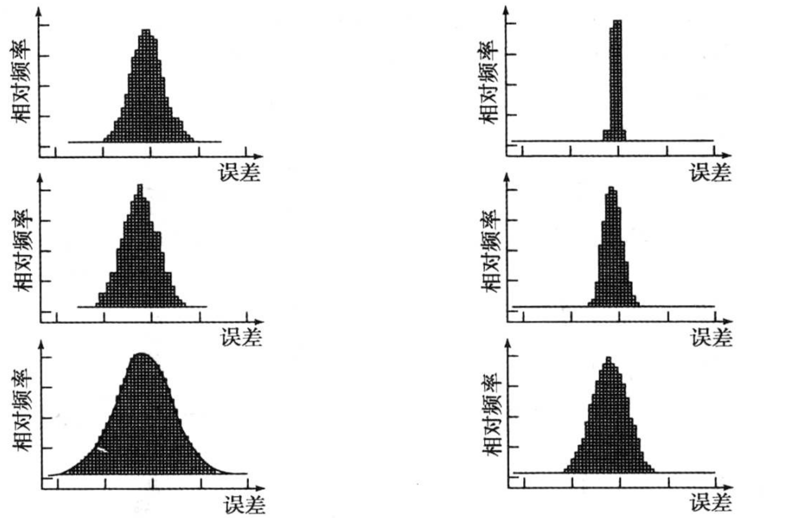
大样本方法不仅能够得到表现优良的估计量；还能告诉我们这些估计量的分布，因而使得我们能够进行假设检验。当一个估计量的干扰在有限样本中太难以确定时，计量经济学家就希望这个估计量的大样本分布能够为有限样本分布提供一个很好的近似。如果这个希望能够实现(经常能够实现)，那么，基于大样本分布的推断就为计量经济学家的推断提供了一个很好的指导，就好像他们知道估计量的有限样本分布一样。

当你自己的数据序列很短时，大样本性质就是不起任何作用的安慰。有些估计量因其大样本性质而流行起来，计量经济学中有一个不断壮大的领域就是对这些估计量的小样本性质进行蒙特卡罗研究。计量经济学家不断发现一些在大样本中表现优异而在小样本中表现糟糕的估计量。

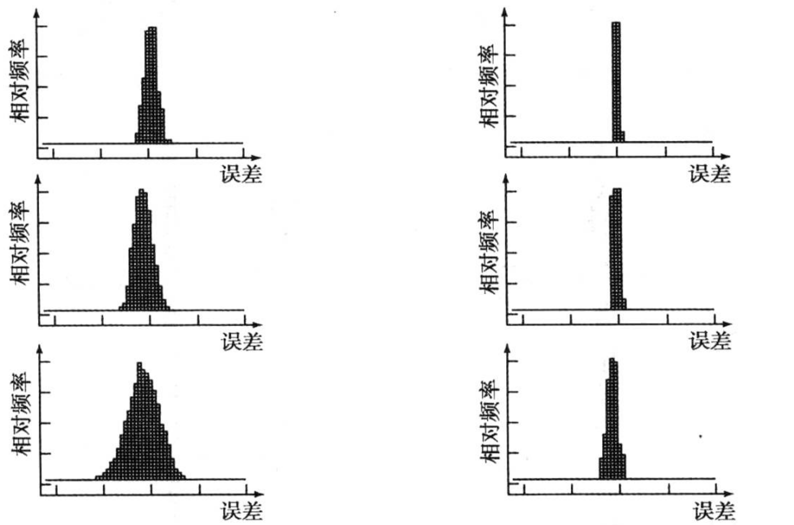
图示一致性和渐近分布

图形能够让我们清晰直观地看到一致性和渐近分布这两个大样本的核心概念。图 12-1 和图 12-2 复制了图 2-6 和图 2-7，它给出了我们最早给出的几个直觉估计量 β_{g1}、β_{g2} 和 β_{g3} 随着数据生成过程中的样本容量从 2 增加到 500 时的蒙特卡罗实验分布。(图 12-1 和图 12-2 中报告的实验是利用蒙特卡罗程序Ⅱ进行的。)随着样本容量的增大，β_{g1}、β_{g2} 和 β_{g3} 的分布都聚集在单个值附近。我们在第 3 章还了解到，在高斯—马尔科夫假定下，β_{g1}、β_{g2} 和 β_{g3} 的分布都有一个期望值 β。因此，这三个估计量聚集在其周围的这单个值就是真实值 β。如果随着样本容量无限增大，这种向 β 的真实值逐渐靠拢的过程无需验证，即如果一个无偏估计量的方差在极限状态缩小至 0，那么，我们就说这个估计量是 β 的一致估计量。

图 12-1 背后的蒙特卡罗实验假定干扰是正态分布的；图 12-2 背后的蒙特卡罗实验假定干扰来自一个高度偏态的离散分布，即以 1/3 的概率取值 -750，以 2/3 的概率取值 375。我们的估计量对如此悬殊的干扰都表现出一致性，这个事实表明，这些估计量的一致性是相当稳健的。到底是什么导致了这个稳健的一致性呢？从数学上讲，β_{g1}、β_{g2} 和 β_{g3} 的一致性都依赖于大数定律。**大数定律**(law of large numbers)说的是，在适当条件下，样本均值的分布在大样本中倾向于聚集在对应的总体均值附近。



a）n=2时 β_{g1}、β_{g2}和 β_{g3}的估计误差　　b）n=20时 β_{g1}、β_{g2}和 β_{g3}的估计误差



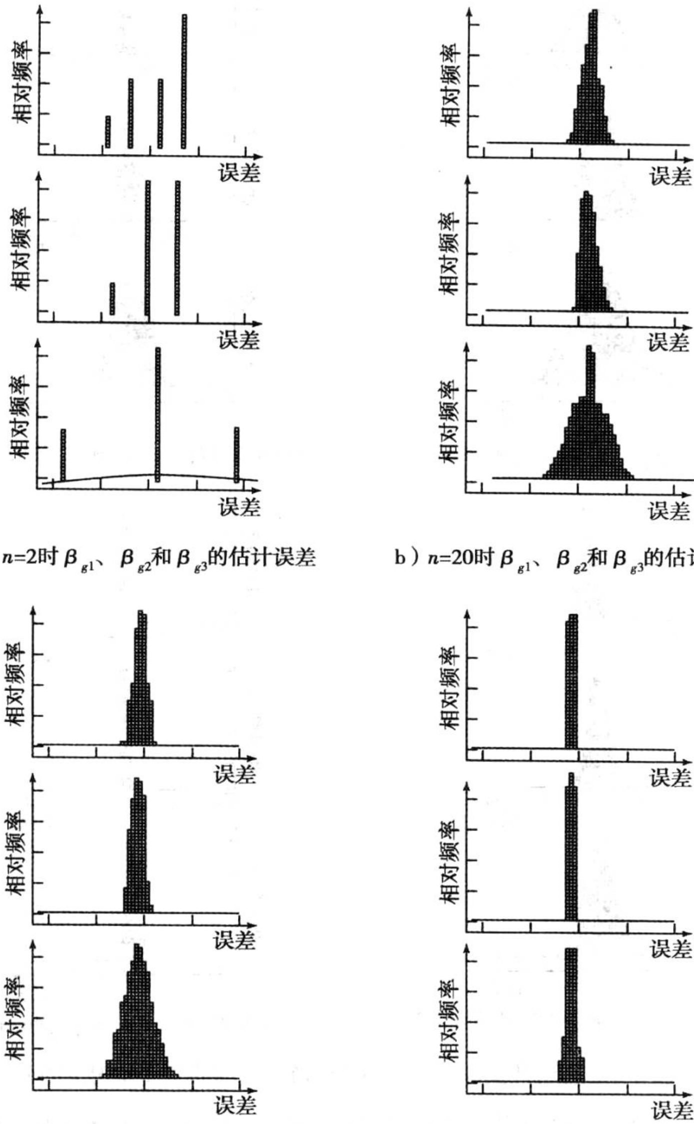
c）n=50时 β_{g1}、β_{g2}和 β_{g3}的估计误差　　d）n=500时 β_{g1}、β_{g2}和 β_{g3}的估计误差

图 12-1　β_{g1}、β_{g2}和β_{g3}在正态干扰和几个不同样本容量情况下的分布

图 12-1 显示出，如果潜在的干扰是正态分布的，那么，对于所有样本容量，β_{g1}、β_{g2}和β_{g3}的分布看起来都很像正态分布——它们应该如此，因为正态分布变量的线性组合本身就是正态分布的。图 12-2 则更加令人吃惊地显示出，在干扰本身相当非正态的情况下，即使对于 20 那么小的样本容量，β_{g1}、β_{g2}和β_{g3}的分布看起来都很像正态分布。尽管干扰是非正态的，但β_{g1}、β_{g2}和β_{g3}在大样本中的分布还是近似正态的，原因来自于统计学上著名的中心（或关键）极限定理。**中心极限定理**（Central Limit Theorem）说的是，在适当条件下，样本均值的分布在大样本中倾向于服从近似正态分布。

当中心极限定理适用于一个估计量时，我们就说这个估计量的渐近分布是正态分布。中心极限定理说明了计量经济学家依赖于正态分布的理由，并说明了即使干扰不是正态分布的，在大样本中使用正态分布进行统计推断的合理性。第 9 章曾提到，就算干扰不是正态分布的，t-检验和 F-检验在大样本中依然适用；这正是中心极限定理保证了 t-统计量和 F-统计量在大样本中近似服从 t-分布

和 F-分布。由于我们通常并不知道模型干扰的实际分布，所以，对于常见的干扰分布，而不仅仅是正态分布，标准的假设检验在大样本中都是可靠的，这令我们倍受鼓舞。



a) n=2时 β_{g1}、β_{g2}和 β_{g3}的估计误差　　b) n=20时 β_{g1}、β_{g2}和 β_{g3}的估计误差

c) n=50时 β_{g1}、β_{g2}和 β_{g3}的估计误差　　d) n=500时 β_{g1}、β_{g2}和 β_{g3}的估计误差

图 12-2　β_{g1}、β_{g2} 和 β_{g3} 在偏态的离散干扰和几个不同样本容量情况下的分布

12.2　渐近无偏性、一致性和概率极限

对诸如无偏性和一致性这类统计性质的直观描述有助于计量经济学家向非专业人员解释他们的工作。但判断哪些估计量具有或不具有某个特定的统计性质，就需要能够用于一些特定估计量的规范定义。本节将定义估计量的两个大样本性质：渐近无偏性和一致性。确定估计量的一致性通常需要我们称之为概率极限的一个新的分析工具——本节也对这个工具进行介绍。

12.2.1　渐近无偏线性

一个自然的大样本性质就是**渐近无偏性**(asymptotic unbiasedness)。如果一个估计量的偏误随着观测次数无限增加具有极限0，那么，这个估计量就是渐近无偏的。一个对每个 n 都无偏的估计量也是渐近无偏的，但许多对每个样本容量 n 都有偏误的估计量却是渐近无偏的。比如，在一个由 n 次观测构成的简单随机样本中，考虑作为总体方差估计量的样本方差。样本方差

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

就是总体方差 σ^2 的一个有偏误的估计量，因为

$$\mathrm{E}(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

但也注意到，随着 n 逐渐变大，$(n-1)/n$ 收敛于1，而样本方差的偏误也随之消失。虽然样本方差是总体方差的一个偏误估计量，但也是总体方差的一个渐近无偏估计量。

12.2.2　一致性

一个理想的估计程序应该在所有时候都是完全正确的。这个崇高的理想可望不可及。不过，如果样本容量足够大，有些估计量几乎总是近似正确的。这种估计量的分布聚集在待估计参数的真实值附近。我们把这样的估计量称为**一致估计量**(consistent estimator)。更准确地讲，如果超过某个样本容量之后，一个估计量的估计值收敛于1的概率就像我们希望真正的参数值那样，那么，这个估计量就是**一致的**(consistent)。不幸但也不可避免的是，即使在大样本中，一致估计量通常仍存在较远偏离我们希望的正确值的较小风险。

一致性的后果之一是，随着样本容量的增大，估计量的分布聚集在待估计参数的真实值附近。图12-3就说明了一个一致估计量 $\tilde{\beta}$ 的分布聚集在真实值 β 附近的情形。该图同时给出了样本容量为5、20、50和500时 $\tilde{\beta}$ 的分布。由于 $\tilde{\beta}$ 是 β 的一致估计量，所以随着样本容量越来越大，估计值的分布就越来越集中，而且这种集中越来越接近 β。

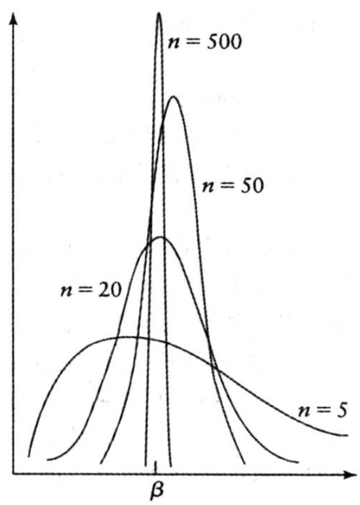


图 12-3　一致估计量的分布随着 n 的增大而逐渐聚集

一个经常遇到的一致估计量的例子就是随着样本容量的增大方差收敛于0的渐近无偏估计量。随着样本容量越来越大，这种估计量的分布聚集在真实值附近。不过，在一些不经常遇到的病态情形中，也有一致估计量不是渐近无偏的情形。一个这种病态的例子，有助于澄清一致性的真正

含义。

假设我们想用一个估计量 $\tilde{\tilde{\beta}}$ 来估计 β，$\tilde{\tilde{\beta}}$ 分别以概率 $1/n$ 和 $(1-1/n)$ 取值 $(\beta+2n)$ 和 $(\beta+1/n)$。这个估计量的病态性在于，随着我们得到的观测越来越多，有得到越来越大的误差 $2n$ 的可能性。这个估计量的期望值是

$$\mathrm{E}(\tilde{\tilde{\beta}})=(\beta+2n)(1/n)+(\beta+1/n)(1-1/n)=2+\beta+1/n-1/n^2$$

通常，它都不等于 β，所以 $\tilde{\tilde{\beta}}$ 是 β 的一个偏误估计量。随着 n 无限增大，$\tilde{\tilde{\beta}}$ 的期望值接近 $\beta+2$，所以，$\tilde{\tilde{\beta}}$ 是一个渐近偏误估计量。

不过，$\tilde{\tilde{\beta}}$ 是 β 的一个一致估计量，因为随着 n 逐渐变大，$\tilde{\tilde{\beta}}$ 处在 β 的一个不断缩小的范围 $(1/n)$ 之内的概率接近于 1。与 $\tilde{\tilde{\beta}}$ 取值 $(\beta+2n)$ 相联系的较大误差越来越不可能出现，越来越有可能出现的 $(\beta+1/n)$ 值则越来越接近于 β。

既不一致又渐近偏误的估计量比一致而又渐近偏误的估计量更加常见。计量经济学家尽一切可能回避这种估计量，但回避这种估计量就需要有识别它们的方法。概率极限就是计量经济学家用来区别一致估计量和不一致估计量的统计工具。

12.2.3 概率极限

计量经济学家认为，一个一致估计量依概率收敛于我们试图得到的真实参数。如果随着样本容量的增大，一个随机变量 β^* 取值像我们希望的那样接近于一个常数 c，那么，我们就说 β^* **依概率收敛于**(converges in probability)常数 c。[2] 在这种情况下，我们把 c 称为 β^* 的**概率极限**(probability limit)或 plim。并不是每个估计量的分布都随着样本容量的增大而逐渐聚集，所以并不是每个估计量都以概率收敛于一个常数。在所有依概率收敛于一个常数的估计量中，只有部分估计量收敛于我们试图得到的真实值，所以只有部分估计量是一致的。我们把如下表述

$$\mathrm{plim}(\beta^*)=d$$

读作“β^* 依概率收敛于 d”，或等价地称之为“β^* 的概率极限是 d”。β^* 是 β 的一致估计量的充分必要条件是

$$\mathrm{plim}(\beta^*)=\beta$$

对于收敛于一个常数的随机变量，一个经常遇到的例子是一个均值为 d 且方差随着样本容量的增大而收敛于 0 的随机变量。这就对应着一个方差随着样本容量的增大而收敛于 0 的无偏估计量的一致性。

计量经济学家常常对某些变量或变量函数的样本均值等随机变量的大样本性质感兴趣，比如干扰的样本方差或干扰与某个解释变量之间的样本协方差。统计学上的大数定律说的是，在适当条件下，样本均值依概率收敛于它在总体中的期望值。计量经济学家通常借助于大数定律将一个随机变量的概率极限与其期望值联系起来。我们已经看到大数定律得以成立的一对条件：①我们所考虑的随机变量是一个样本均值，它的期望值总等于某个有限数值 d；②样本均值的方差随着样本容量的增大而收敛于 0。这个定律在很多其他环境中也成立，但并不是在所有情形中都成立。无界方差常常使得大数定律不能成立。附录 12A 对概率极限和一致性有更规范的讨论。

12.3 $\boldsymbol{\beta}_{g4}$ 和 $\hat{\boldsymbol{\beta}}_1$ 的一致性

我们在第 3 章和第 4 章已经了解到，在高斯—马尔科夫假定下，β_{g4} 和 $\hat{\beta}_1$ 是最优线性无偏估计

量。这些性质已经足够有吸引力，所以就没有理由仅仅为了找到最优二乘估计量的一个更有吸引力的性质而使用概率极限的工具。一致性和概率极限的现实价值在于分析比高斯—马尔科夫框架更加复杂的数据生成过程。不过，既然我们对最小二乘估计量和高斯—马尔科夫假定如此熟悉，那它们就能为我们研究一致性和概率极限提供富有启发意义的直观印象。因此，本节将在高斯—马尔科夫假定下分析 β_{g4} 和 $\hat{\beta}_1$ 的一致性。

虽然普通最小二乘估计量 β_{g4} 和 $\hat{\beta}_1$ 在高斯—马尔科夫假定下是一致的，但我们还是必须对我们在样本容量不断扩大时所观察到的其他 X 值做进一步的假定。本节在高斯—马尔科夫假定下对 β_{g4} 的一致性的证明依赖于大数定律的一个基本应用。本节对 $\hat{\beta}_1$ 的一致性的证明更加详尽；它成为以后章节中分析其他估计量的一致性的一个模板。

12.3.1　高斯—马尔科夫假定下 $\boldsymbol{\beta_{g4}}$ 的一致性

我们在第 3 章了解到，在高斯—马尔科夫假定下，过原点直线斜率的最优线性无偏估计量是

$$\beta_{g4}=\frac{\sum X_iY_i}{\sum X_i^2}$$

它的期望值等于 β，方差等于 $\sigma^2/\sum X_i^2$。由于 β_{g4} 是一个无偏估计量（$\mathrm{E}(\beta_{g4})=\beta$），所以，如果它的方差随着样本容量的不断扩大而收敛于 0，那 β_{g4} 就是一致的。因此，只要 $\sum X_i^2$ 随着样本容量的扩大而无限变大，β_{g4} 就是一致的。

但高斯—马尔科夫假定并没有告诉我们，在样本容量不断增大时，在分析中将增加哪些 X 值。我们只知道，在每个样本中，如果有第 i 次观测的话，第 i 次观测的 X 值就是 X_i。如果我们要保证 β_{g4} 的方差在大样本中收敛于 0，我们还需要假定 $\sum X_i^2$ 随着 n 的不断增大而无限递增。这是一个不算过分的要求，但它还是排除了一些数据生成过程。比如，一个利用渡渡鸟（Dodo bird，现已灭绝）的数量作为解释变量的时间序列研究就不满足 $\sum X_i^2$ 随着 n 的不断增大而无限递增的要求——在渡渡鸟灭绝以后的时期，这个和就不再变大；于是此后的 X_i 都等于 0。类似地，那些常常能够得到满足的假定与高斯—马尔科夫假定一起就使得我们的其他无偏估计量成为一致估计量。

注意到，β_{g4} 的一致性还取决于它是 σ^2 有限的。如果在我们的数据生成过程中，干扰具有无限方差，那么，估计量方差的分子中 σ^2 一项就可能抵消分母不断变大的影响，方差就有可能不是收敛于 0。在经济应用中，我们需要检查有限方差的假定是否能够得到保证。在横截面数据中，方差几乎总是有限的。但在时间序列数据中，方差有时会无限增大。无界方差使得大数定律不能适用，因而普通最小二乘法就是不一致的。

12.3.2　计算概率极限

计算概率极限有如下常见法则：

(1) $\mathrm{plim}(c)=c$，其中 c 是一个常数。

(2) 常数序列的概率极限等于常数序列的极限。

若 α 和 β 是随机变量，而且 $\mathrm{plim}(\alpha)$ 和 $\mathrm{plim}(\beta)$ 是有限的，则：

(3) $\mathrm{plim}(\alpha+\beta)=\mathrm{plim}(\alpha)+\mathrm{plim}(\beta)$。

(4) $\mathrm{plim}(\alpha\beta)=\mathrm{plim}(\alpha)\mathrm{plim}(\beta)$。

(5) 若 $\mathrm{plim}(\alpha)\neq 0$，则 $\mathrm{plim}(\beta/\alpha)=\mathrm{plim}(\beta)/\mathrm{plim}(\alpha)$。

(6) 对于连续函数 g，有 $\mathrm{plim}(g(\beta))=g(\mathrm{plim}(\beta))$。

（大数定律说的是，在适当条件下，样本均值与其期望值之差依概率收敛于 0。如果样本均值的期望值是有限的，而且它的方差随着 n 的无限增大而具有极限 0，那么，这个定律都是可以照常使用的。）这些关系式的明显特征是它们既不要求 α 和 β 相互独立，又不要求 g 是线性函数，这就使得

乘积或比率的概率极限比期望值更容易计算。

这些运算法则类似于期望运算法则，但也有一些重要差别。第一，期望适用的随机变量在我们所看到的不同样本中有不同的实现值，而概率极限适用的随机变量在我们连续抽取越来越大的样本时出现。第二，除非我们所考虑的随机变量在统计上是独立的，否则，概率极限运算法则(4)、(5)和(6)一般在期望运算中都不成立，但无论随机变量之间的相关性如何，它们都适用于概率极限。正是因为预算法则(4)、(5)和(6)对统计上独立和不独立的随机变量都适用，才使得对一致性的分析比对偏误的分析更加容易。

12.3.3 $\hat{\beta}_1$ 的一致性

我们已经了解到，在高斯—马尔科夫假定下，只要 $\sum X_i^2$ 随着样本容量的扩大而无限变大，普通最小二乘法就能一致地估计过原点直线的斜率。这里，我们在高斯—马尔科夫假定下，利用上一节的概率极限运算法则来证明 $\hat{\beta}_1$ 的一致性。我们再次需要对 X 随着 n 的增大而新出现的取值做一个额外假定。这个假定就是 $\lim(\sum x_i^2/n)=Q$ 是一个不等于 0 的常数。就像对 β_{g4} 的一致性的证明那样，证明 $\hat{\beta}_1$ 的一致性就是要考查 $\hat{\beta}_1$ 的均值和方差。不过，对 $\hat{\beta}_1$ 的一致性多少更加详尽的证明，能够为分析除最小二乘估计量之外其他估计量的一致性提供一个模板，所以本节将给出一个较为详尽的证明。

记得

$$\hat{\beta}_1=\frac{\sum x_iY_i}{\sum x_i^2}=\beta_1+\frac{\sum x_i\varepsilon_i}{\sum x_i^2}$$

我们可以把它重新写成

$$\hat{\beta}_1=\beta_1+\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}$$

如果它依概率收敛于 β_1，也就是说，如果

$$\operatorname{plim}(\hat{\beta}_1)=\operatorname{plim}\left(\beta_1+\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}\right)=\beta_1$$

那么，最小二乘估计量就是一致的。概率极限的第 3 个运算法则是，和的概率极限等于概率极限之和。因此

$$\operatorname{plim}\left(\beta_1+\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}\right)=\operatorname{plim}(\beta_1)+\operatorname{plim}\left(\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}\right)$$

第一个法则说的是，一个常数的概率极限等于这个常数。因此，

$$\operatorname{plim}(\beta_1)+\operatorname{plim}\left(\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}\right)=\beta_1+\operatorname{plim}\left(\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}\right)$$

第 5 个法则说的是，如果一个比率的分母的概率极限不等于 0，那么这个比率的概率极限就等于分子和分母的概率极限之比。因此，如果 $\operatorname{plim}(\sum x_i^2/n)$ 不等于 0，那么

$$\beta_1+\operatorname{plim}\left(\frac{\sum x_i\varepsilon_i/n}{\sum x_i^2/n}\right)=\beta_1+\frac{\operatorname{plim}(\sum x_i\varepsilon_i/n)}{\operatorname{plim}(\sum x_i^2/n)} \tag{12-1}$$

样本容量的扩大需要增加新的 X 值，这些新的 X 值可能会改变 X 的均方差。[⊖]因此，在样本容量不断变化时，最好把 X 的均方差看成是一系列可以变化的数字。我们新增加的假定是，这个序列的极限是 Q。第 2 个运算法则说的是，一个常数序列的概率极限就等于该序列的极限，所以 $\operatorname{plim}(\sum x_i^2/n)$ 就等于 $\lim(\sum x_i^2/n)$，根据假定，这个极限等于 Q。我们可以把式(12-1)重新写成

⊖ 原书使用 X 的平方的均值，与公式不符。——译者注

$$\beta_1+\frac{\operatorname{plim}\left(\frac{1}{n}\sum x_i\varepsilon_i\right)}{\operatorname{plim}\left(\frac{1}{n}\sum x_i^2\right)}=\beta_1+\frac{\operatorname{plim}(\sum x_i\varepsilon_i/n)}{Q}$$

因此，$\hat{\beta}_1$ 的一致性取决于 $\operatorname{plim}(\sum x_i\varepsilon_i/n)$ 是否等于 0。

高斯—马尔科夫假定意味着 $\operatorname{plim}(\sum x_i\varepsilon_i/n)$ 等于 0。因此，如果 $(\sum x_i\varepsilon_i/n)$ 的方差随着 n 的增大而收敛于 0，那么 $\operatorname{plim}(\sum x_i\varepsilon_i/n)$ 实际上确实等于 0，因为它的分布将随着样本容量的扩大而逐渐聚集在其均值上。$(\sum x_i\varepsilon_i/n)$ 的方差是多少呢？由于在高斯—马尔科夫假定下，x 和 n 都是常数，而且干扰之间的协方差都等于 0，所以这个方差就是

$$\operatorname{var}\left(\frac{1}{n}\sum x_i\varepsilon_i\right)=\frac{1}{n^2}\sum x_i^2\operatorname{var}(\varepsilon_i)$$

因此，我们发现

$$\frac{1}{n^2}\sum x_i^2\operatorname{var}(\varepsilon_i)=\frac{1}{n^2}\sum x_i^2\sigma^2=\left(\frac{\sigma^2}{n}\right)\left(\frac{1}{n}\sum x_i^2\right)$$

随着 n 的增大，它的极限就是

$$\lim[(\sigma^2/n)\sum x_i^2/n]=\lim(\sigma^2/n)\lim(\sum x_i^2/n)=0Q=0$$

普通最小二乘估计量 $\hat{\beta}_1$ 就是 β_1 的一致估计量。在本书网站(www.aw-bc.com/murray)上的补充材料中，附录 12C 在多元回归情形中用矩阵代数证明了普通最小二乘估计量的一致性；在那里，也需要对新增加的 X 施加一些不太过分的假定。这些不太过分的假定既保证了解释变量在大样本中不会出现完全共线性，也保证了 X 值不会遇到渡渡鸟那样的问题。

12.3.4　一致估计量和数据生成过程的设定

对 $\hat{\beta}_1$ 的一致性的上述分析，具有计量经济研究中经常遇到的一个特征。计量经济学家并非总是假定他们所需要做出的最起码假定，他们所做的假定常常比严格必需的假定更强。在一定程度上，式(12-1)就告诉了我们 $\hat{\beta}_1$ 要成为 β_1 的一致估计量所需要的条件，就需要

$$\operatorname{plim}(\sum x_i\varepsilon_i/n)=0 \qquad (12\text{-}2)$$

$$\operatorname{plim}(\sum x_i^2/n)=Q\neq 0 \qquad (12\text{-}3)$$

在式(12-2)不成立但(12-3)成立的时候，普通最小二乘估计量既是渐近偏误的，又是不一致的；估计量的分布聚集在一个错误的数值附近。

但计量经济学家在证明 $\hat{\beta}_1$ 的一致性时，不是直接假定式(12-2)和式(12-3)，而是利用大数定律，从如下扩展之后的高斯—马尔科夫假定来推导式(12-2)和式(12-3)：

- 解释变量在不同样本中保持不变。
- $E(x_i\varepsilon_i)=0$ (12-4)
- $\lim\limits_{n\to\infty}(\sum x_i^2/n)=Q\neq 0$ (12-5)

当解释变量在不同样本中保持不变时，式(12-4)和式(12-5)是比式(12-2)和式(12-3)成立所必需的假定更强。为什么计量经济学家总是做比需要更强的假定呢？他们为什么不就假定式(12-2)和(12-3)是正确的呢？

假定式(12-4)和式(12-5)而不是假定式(12-2)和式(12-3)有一个优点：对假设 $E(x_i\varepsilon_i)=0$ 可以直接进行检验，但我们没有办法实证地分析 $\operatorname{plim}(\sum x_i\varepsilon_i/n)$，因为后者是一个无限序列的极限。计量经济学家在设定模型时的目标之一就是要让他们的假定可以检验。计量经济学家经常使用对期望的假定，并结合大数定律来证明渐近收敛结论，其第二个原因是从经济理论中得到对期望值的约束条件比得到对概率极限的约束条件更加常见。因此，把数据生成过程植根于经济理论的计量经济学家更喜欢使用从经济理论中推导出来的期望值假定，并结合大数定律来判断其估计量的一致性。