21世纪高等院校金融学教材新系

Financial Econometrics

张成思 著

(中国人民大学)

金融计量学

——时间序列分析视角

东北财经大学出版社

Dongbei University of Finance & Economics Press

21世纪高等院校金融学教材新系

Financial Econometrics

张 成 思 著

(中国人民大学)

金融计量学

——时间序列分析视角

东北财经大学出版社
Dongbei University of Finance & Economics Press

图书在版编目（CIP）数据

金融计量学——时间序列分析视角 / 张成思著.—大连：东北财经大学出版社，2008.7
（21 世纪高等院校金融学教材新系）
ISBN 978 - 7 - 81122 - 273 - 9

Ⅰ.金… Ⅱ.张… Ⅲ.金融 - 计量经济学 - 高等学校 - 教材 Ⅳ.F830

中国版本图书馆 CIP 数据核字（2008）第 056288 号

东北财经大学出版社出版
（大连市黑石礁尖山街 217 号 邮政编码 116025）
总 编 室：（0411）84710523
营 销 部：（0411）84710711
网 址：http: //www.dufep.cn
读者信箱：dufep @ dufe.edu.cn
大连北方博信印刷包装有限公司印刷 东北财经大学出版社发行

幅面尺寸：170mm×240mm 字数：413 千字 印张：16 1/4
2008 年 7 月第 1 版 2008 年 7 月第 1 次印刷

责任编辑：田玉海 责任校对：惠恩乐 何 群
封面设计：沈 冰 版式设计：钟福建

ISBN 978 - 7 - 81122 - 273 - 9
定价：28.00 元

前 言

金融学是社会科学中最具有魅力的学科之一，因为无论是微观层面的公司金融（corporate finance），还是宏观层面的货币金融、国际金融等，都与人们的经济生活紧密相关。而随着现代金融学科的快速发展，尤其是以威廉·夏普（William Sharpe）、默顿·米勒（Merton Miller）、哈里·马科维茨(Harry Markowitz)以及罗伯特·默顿（Robert Merton）在布莱克—斯科尔斯（Black-Scholes）期权定价公式方面的不懈研究，以及罗伯特·恩格尔(Robert Engle)和克莱夫·格兰杰(Clive Granger)在研究金融资产收益波动性（volatility）聚类现象的时序分析方面的贡献，使得金融计量方法不断推陈出新，而现代金融学的教学和研究也日益突出计量分析的重要地位。

金融计量学（Financial Econometrics）正是在这个大的背景下发展起来的。从名称不难看出，金融计量学实质上是运用经济计量方法研究金融领域的相关问题。在金融计量学发展的早期，尤其是 20 世纪 90 年代之前，金融计量学一般被划归为计量经济学的一个分支。然而，随着金融学的独立发展和学科体系的日益完善，用来研究金融学的现代计量方法，在 20 世纪 90 年代之后也随之出现了一番波澜壮阔的发展景象，特别是金融时间序列分析方法在金融领域得到了广泛的应用和发展。

近年来，随着中国的金融学科的快速发展和学科体系的逐步成熟，现代金融计量方法也越来越多地被用来分析现实中的金融问题。当前，“金融计量学”等相关课程也逐渐成为国内许多高校为金融学、经济学、工商管理以及财政学等相关专业的本科生和研究生开设的专业必修课。

很自然地，选择合适的教材就成为教师教学和学生学习的一个关键环节。当然，国外近年来不乏相关的优秀教材，可以直接选用。但是，对于初次接触金融计量学方面知识的学生，选用外版教材可能存在下面的问题：一是如果直接使用英文原版书籍，很多术语和重要的知识点可能不容易理解；二是如果使用引进的翻译版，由于各方面原因，可能导致译稿与原版内容之间存在信息传递的缺失，尤其是一些金融计量学中特有的专业术语和含义，转译过程中经常出现漏译和误译现象，导致某些重要的专业知识信息的损失。因此，广大教师和学生期望能够得到难度适宜的中文本版金融计量经济学书籍，作为案头工具书和学习参考资料。

有鉴于此，作者结合在英国曼彻斯特大学经济学院和中国人民大学财政金融学院的国外和国内教学经验，撰写了《金融计量学——时间序列分析视角》一书，旨在作为符合中国学生阅读习惯，又能帮助读者领略国际最新知识动态的本版金融计量学教材。

本书的内容涵盖金融计量的主要分析方法，同时强调理论与实际应用相结合，并且突出本版教材的特点，应用的实际例子大部分以中国的经济金融数据为主，并包括非线性金融时间序列分析方法等金融计量前沿知识。为提高本书的可读性，作者将涉及到的比较繁难的内容以简单浅显的语言形式和容易理解的图表形式解读出来，并且结合金融计量软件讲解一些具体数据处理和回归操作过程，形式新颖，期望使读者阅而不烦。

本书适合用作金融学、经济学和工商管理等专业的研究生或高年级本科生教材，对于具有计量经济学基础或者正在学习计量经济学基础课程的学生，本书不失为一本很好的学习用书。另外，对于具有一定金融学或经济学基础的从业人员和科研工作者，本书也可以

作为一本案头参考书。当然，对于以前没有计量经济学基础的读者，建议先学习一定的基础知识，如参考古扎拉蒂的第 4 版《计量经济学》（Damodar Gujaradi, *Basic Econometrics*，2003）或者伍德里奇的《计量经济学导论——现代观点》（Jeffrey Wooldridge，*Introductory Econometrics—A Modern Approach*，2003），再来学习本书，效果可能会更好。

为方便广大教师在选用本教材时能够根据具体的课程规划和课时安排有针对性地使用，下面就本书的内容设计、结构和各部分的关联加以说明。本书共分为 11 章，第 1 章是金融计量学的概括介绍，包括现代金融计量软件的初步介绍，尤其是 EViews 软件的使用。第 2 章介绍差分方程和滞后运算法。第 3、4、5、6 章分别介绍平稳金融时间序列和非平稳金融时间序列，包括 AR、MA、ARMA 以及非平稳时间序列的检验方法。这 4 章是金融计量经济学的基础与核心之一，教师备课过程中可以适当详细地讲解本部分内容。第 7 章是对向量自回归模型（VAR）的介绍。第 8 章进一步介绍结构性向量自回归模型（SVAR），并说明相关模型的应用思路。第 9 章对协整和误差修正模型进行系统介绍，包括著名的 Engle-Granger 协整分析方法和以向量模型为基础的 Johansen 协整分析方法。第 10 章介绍金融时间序列分析中的条件异方差模型，即 ARCH 和 GARCH 模型。这两章分别运用汇率、通胀率等数据，说明建立模型与实证检验的具体步骤，如果在课时有限的情况下，可以略去实证部分的内容。本书第 11 章介绍金融计量经济学中的非线性时间序列模型，包括马尔可夫区制转移模型和门限模型等，建议在课时充分的条件下，有重点地详细讲解。最后，附录 A1 回顾了金融计量学的背景知识，即矩阵代数和经典线性回归模型。这一部分被刻意安排在正文后，主要是考虑到读者背景知识的差别，例如，已经具有计量经济学基础的同学，可以直接阅读本书第 1 章到第 11 章关于金融计量学的核心内容。而对于需要补充背景知识的读者，则可以先学习一下附录 A1 的内容，为理解前 11 章的内容做好铺垫。附录 A2 为金融计量学常用的统计表。

在本书的撰写过程中，我得到了中国人民大学财政金融学院、英国曼彻斯特大学经济学院和香港中文大学的朋友和同事的大力支持和帮助，在此向他们表示感谢。同时，在著书过程中，作者还得益于与诸多世界知名计量经济学家和同行的交流与探讨，如哈佛大学的 James Stock 教授和 Daniel Ahn 博士、普林斯顿大学的 Mark Watson、麦吉尔大学的 Russell Davidson、纽约大学的 Bai Jushan、耶鲁大学的 Peter Phillips、波士顿大学的 Pierre Perron、威斯康星大学的 Bruce Hansen、诺丁汉大学的 Paul Mizen、华威大学的 Martin Ellins、布里斯托大学的 Joel Clovis、香港科技大学的 In Choi 以及曼彻斯特大学的 Denise Osborn、Alastair Hall、Chris Orme、Ralf Becker 和 Simon Peters，在此对这些直接或间接帮助过我的人表示感谢。最后，我还要感谢我在曼彻斯特大学和中国人民大学财政金融学院的学生，特别是 Robert Enderson、Simon Watts、刘尚、边慧、李晓刚、朱哲、成功、张捷、刘洋、张延琪、唐诗瑶等，他们对本书部分内容曾经提出的问题和建议帮助我在本书的终稿当中尽量少地出现疏漏。

本书的出版受到 2007 年度国家级双语教学示范课程建设项目基金支持，作者为该项目“金融计量学”的主持人，对该项目基金的资助特别表示感谢！

虽然在撰写本书的过程中，作者倾注了大量精力和热情，认真审阅每一部分内容，但是由于自身学识、能力和思考问题的视角可能存在的限制，书中不足之处在所难免，真诚地希望广大读者和学界同仁不吝批评指正，或者通过电子邮件 zhangchengsi@yahoo.com.cn 反馈您的意见。

张成思

2008 年 6 月于北京

目 录

第1章

金融计量学介绍

★ 导读

金融时间序列变量是现代金融学中实证研究的核心。债券、股票、期货、期权等微观金融工具的收益率和风险等都涉及到金融时间序列分析。而宏观金融层面的利率、汇率和通货膨胀率等变量更是因时而动，有着与横截面数据(cross-section data)截然不同的特点。本章从基本的金融时间序列入手，回顾了金融学建模（financial modeling）过程中涉及的基本概念，如股票债券等的收益率计算公式等。同时，本章还介绍了后续内容经常用到的相关统计量和相关的统计知识，包括随机变量和随机变量的各阶矩（moment）等重要概念。最后，本章对金融时间序列分析常用的计量软件做了入门式的介绍，以期使读者更快地掌握应用理论处理实际问题的能力。

作为开篇介绍，本章主要介绍金融计量学的背景知识，为后续从事实证金融计量分析奠定基础。有人认为："金融计量学研究的就是时间序列的资产价格评估理论与实践。"这种诠释金融计量学的角度只能说涵盖了该学科领域的部分内容。事实上，金融计量学涵盖的内容要远比这种描述丰富得多。例如，现代金融计量学家所面临的问题不仅仅是资本资产定价模型（CAPM）的应用，也涉及到更广泛的金融风险计量计算、股价波动性问题分析、汇率决定机制的计量分析，以及通货膨胀动态路径的预测分析等。

因此，本书的内容没有局限在资本资产定价模型或者行为金融领域常用的事件分析方法上面，而是以**金融时间序列分析**为主线，介绍金融领域常用的一维和多维计量模型。本书不仅介绍现代金融计量方法中的理论模型，比如随机游走模型、金融时序波动模型和协整模型等，而且更注重理论模型的实际应用讲解。这样，我们希望本书能为金融学、经济学专业的学生和金融领域的研究人员提供理解并从事金融计量分析的必要技术背景。

虽然本书对资本资产定价模型和事件分析等金融市场计量模型也有所涉及，但重点放在相关理论的金融时间序列分析方法和应用上面，而关于 CAPM 的理论内容以及事件研究的非金融时间序列分析方法，建议读者参阅坎贝尔等人（1997）的专著《金融市场计量经济学》，以便最大化读者的读书、获知的效用，也避免资源的浪费。

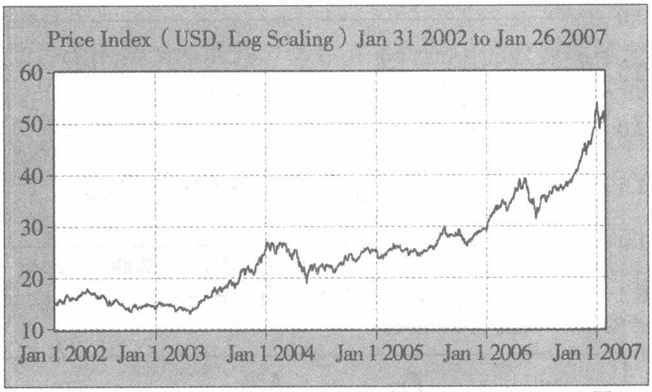
为给读者以更好的启发，本章的第 1 小节将简要介绍金融时间序列的基本概念，初步识别金融时序变量的图示特征；第 2 小节介绍金融计量学中常用的概念和相关统计学知识。在本章的最后一个小节，我们简单介绍了常用的金融计量软件，WinRATs, Matlab 以及 SAS 等，并详细介绍了现代金融时序分析领域最为流行的软件之一 EViews 5.0 以上版本的使用方法。

§1.1 金融时间序列的定义与实例

广义地讲，将某种金融随机变量按出现时间的先后顺序排列起来称为金融时间序列。从现实世界的角度看，金融时间序列就是指在一定时期内按时间先后顺序排列的金融随机变量。如果我们暂时不去深究这里提到的"随机"概念（随机变量的概念将在下一小节中阐述），并假定变量的排序总是按时间先后排列的，那么金融时间序列最显著的特征就是其与"时间"紧密相联。一般来说，金融时间序列变量，有时也简称为金融时序变量，由两个明显的要素组成，即时间跨度和序列的频率。

例如，图 1—1(a)刻画了摩根士丹利国际资本集团(Morgan Stanley Capital International)公布的中国国际股票价格指数 1992 年 12 月至 2006 年 12 月期间的月度数据。显然，该股票价格指数的时间跨度为 1992—2006 年，频率为月度。而图 1—1（b）描绘了中国国际股票价格指数 2002 年 1 月 31 日至 2007 年 1 月 26 日期间的日度数据，这一时序变量的时间跨度为 2002—2007 年，频率为日度。

（a）1992 年 12 月—2006 年 12 月



（b）2002 年 1 月 31 日—2007 年 1 月 26 日（5 天/周）

图 1—1　中国国际股票价格指数

数据来源：Morgan Stanley Capital International。

为进一步了解不同金融时间序列数据的特点，深入理解金融时间序列定义，图 1—2 到图 1—5 分别描绘了人民币/美元汇率、美元/英镑汇率、中国消费者价格指数（CPI）通胀率以及中国货币总量（M1）的增长率时间序列。

从这几幅图可以看到，不同的金融时间序列变量展示出各种各样的变动轨迹，经济学者经常把金融时间序列变量的这种随时间变化的轨迹称为“动态路径”，其中“动态”一词的含义实质上就是指“随时间变化”。我们知道，在 2005 年 7 月之前，人民币汇率政策施行的是钉住美元的政策（pegged exchange rate），如果将 2000—2004 年的人民币/美元汇率的时间序列图示画出来，那么我们观察到的将是一条近乎平坦的水平线。而 2005 年 7 月以后，出于各方面因素考虑，我国开始施行有管理的浮动汇率制度，因此，我们可以从图 1—2 中清楚地看到人民币不断升值的趋势。由于图 1—2 对应的是采用直接报价法（1 美元兑换的人民币数值）的汇率，所以我们观察到了一条向下倾斜的曲线，反映的是人民币不断升值的动态路径。

而同样是汇率的时间序列图示，图 1—3 描绘的美元/英镑的汇率时序图展示的却是一幅截然不同的动态趋势。简单地说，我们看到的美元/英镑的汇率走势呈现上下波动的态势，而不是单一的下降或上升趋势。我们将在本书的第 3~5 章学习到，这两种不同的时序图表现的是具有不同统计特性的时间序列变量，在金融计量学中通常使用“平稳时间序列”和“非平稳时间序列”来区分它们的本质属性。

对于图 1—4 和图 1—5 描绘的分别是我国 CPI 通膨胀率和货币总量（M1）增长率情况。不难看出，这两个金融时间序列变量的动态走势也表现出相当大的差异。所以，似乎很难找到某一种金融计量模型，能够完全捕捉或刻画各种不同的金融时间序列变量的不同特征。

图 1—2　人民币/美元汇率：2005 年 7 月 1 日—2007 年 1 月 12 日

数据来源：中国人民银行。

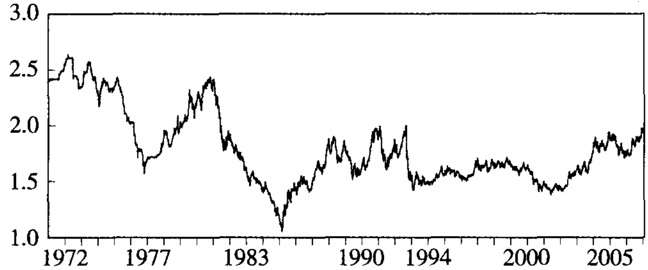


图 1—3 美元/英镑汇率：1971 年 1 月 4 日—2007 年 1 月 12 日

数据来源：Federal Reserve Bank of St. Louis。

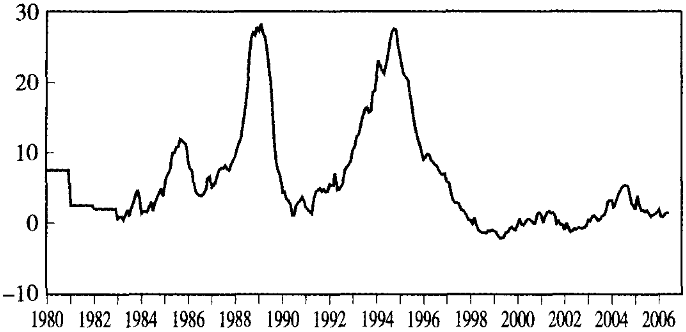


图 1—4 中国 CPI 通胀率：1980 年 1 月—2006 年 6 月

数据来源：中国国家统计局、《中国经济景气月报》。

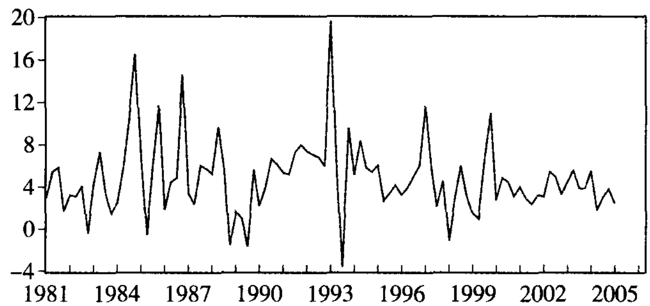


图 1—5 中国 M1 增长率：1981 年第一季度—2005 年第一季度

数据来源：中国人民银行。

§1.2 金融时间序列分析中的基本概念

1.2.1 增长率与收益率

在金融时间序列分析中，我们经常会遇到“价格”的概念，例如，股票价格、债券价格、资产价格、人民币的美元价格以及和我们生活休戚相关的商品价格等。然而，金融计量学家可能更加关注的是价格的增长率或者是金融产品的收益率。之所以如此，最重要的一条原因是增长率要比价格变量具有更吸引人的统计特性。如果我们使用更为专业和更正式的语言来明确解释这种统计特性，那么就是“价格的时间序列一般都含有时间趋势（time trend）成分，而其增长率通常为平稳时间序列（stationary time series）”。需要指出的是，如果读者现在对金融时间序列分析的理论知识还了解得不多，不必被不熟悉的定义和概念所困扰，随着本书各章的循序进展，相信很快就会习惯各种金融计量学中的

所谓的行话（jargons）。

我们再回到上面讨论的增长率概念，即在许多情况下的收益率概念。为方便说明，使用P_t代表某一金融产品在时刻 t 的价格，同时假定与该金融产品相关的红利忽略不计，那么该金融产品的简单净收益率（simple net return）R_t定义为：

$$R_t = 100\% \times \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1.1}$$

公式（1.1）是大家熟悉的增长率计算公式，而在金融时间序列分析中经常会用到另外一个相关的但不完全相同的概念，即连续复合收益率（continuous compounding return），对于单期连续复合收益率，定义为：

$$r_t = 100\% \times \ln(\frac{P_t}{P_{t-1}}) \tag{1.2}$$

其中：ln 代表自然对数运算符号。对于多期（multi-period）来说，比如 k 期, 连续复合收益率定义为：

$$r_t(k) = \ln\left[(\frac{P_t}{P_{t-1}})(\frac{P_{t-1}}{P_{t-2}})\cdots(\frac{P_{t-k+1}}{P_{t-k}})\right] = r_t + r_{t-1} + \cdots + r_{t-k+1} \tag{1.3}$$

所以，多期连续复合收益率是单期的加总。在金融计量分析中，使用连续复合收益率不仅简化了数学计算（即把乘积形式转化为求和形式），而且简化了收益率统计特性的计量建模分析过程。有关这方面更多的阐释，可以参阅 *The Econometrics of Financial Markets*（Campbell *et al.* 1997）。

同时，在处理实际问题的过程中，我们经常需要把各种不同频率的增长率转化为具有统一可比性的增长率或收益率形式，而"年度化增长率"（annualized growth rate）便是最常见的概念之一。年度化增长率反映的是在单期增长率不变的情况下，一个变量在一年的跨度期间内发生变化的幅度。近年来，随着中国金融市场的不断发展和完善以及资本市场的迅速成长，"年度化"一词在各大银行营业厅的金融产品如基金、股票等的宣传材料中屡见不鲜。可见，这一概念已经不仅仅是出现在金融计量学教材中的学术名词了，而是已经与人们的生活联系得相当密切了。

那么，如何计算不同频率金融时间序列的年度化增长率或收益率呢？这与研究的数据频率紧密相关。例如，对于季度频率数据，年度化的增长率计算公式可以近似地写成：

$$100\% \times \ln(\frac{P_t}{P_{t-1}})^4 = 400\% \times \ln(\frac{P_t}{P_{t-1}}) \tag{1.4}$$

依此类推，对于月度频率的数据，年度化的增长率计算公式是：

$$100\% \times \ln(\frac{P_t}{P_{t-1}})^{12} = 1\,200\% \times \ln(\frac{P_t}{P_{t-1}}) \tag{1.5}$$

需要注意的是，上面介绍的年度化增长率计算公式从严格意义上讲是一种近似计算，对应的严格意义上的计算公式应该分别是：

$$100\% \times [(\frac{P_t}{P_{t-1}})^4 - 1]$$

和

$$100\% \times [(\frac{P_t}{P_{t-1}})^{12} - 1]$$

但是根据基本的数学知识我们知道，在一般情况下，这两种算法给出的值非常相近。由于自然对数在金融计量和经济计量中的应用方便而快捷，所以计量学家经常使用前者来

计算金融时间序列近似的年度化增长率。在本书中如果未作特别说明，**年度化增长率均使用含有自然对数的计算公式计算。**

1.2.2 随机变量与随机过程

如果从统计学角度来定义随机变量，我们可以从概率（欧几里德）空间的定义出发，利用波莱尔集合（Borel Set）对随机变量进行定义。例如，Spanos（1986，2000）在其经典的初级统计学教材中，对随机变量给出了严格的统计定义。当然，在其他高级数理和概率理论教材，如 Shiryaev（1996）在金融数学领域很著名的教材等，都给出了随机变量以及随机过程的复杂定义。

但是，我们这里避免让读者陷入繁难晦涩的定义中，而是从金融计量学角度出发采用更为直观的方式来对随机变量和随机过程进行定义。我们可以考虑一个非常简单的经典一元线性回归方程：

$$y_t = c + \beta x_t + \varepsilon_t,\ \ \varepsilon_t \sim N(0,\sigma^2) \tag{1.6}$$

其中：$\varepsilon_t \sim N(0,\sigma^2)$表示$\varepsilon_t$服从均值为 0、方差为$\sigma^2$的正态分布。注意，在很多教材中，经常把正态分布也称为高斯分布（Gaussian distribution），读者在阅读不同资料时应注意二者之间的等同性。

从以前的学习中我们知道，在回归方程（1.6）中，误差项ε_t就是一个随机变量，这里假设这一随机误差变量服从正态分布。在更多的情形下，随机变量ε_t被假设服从**独立一致性分布**（independently and identically distributed），或者简记做 *i.i.d.*。

对于 *i.i.d.*假设，其含义是，无论随机项ε_t是否服从正态分布，在各个时刻点该随机变量服从完全一致的分布特征，并且彼此互相独立。

在上面用的例子中，我们刻意使用了下标 t 来标识时间序列变量 y 和 x。因为 y_t 序列是随机误差项的线性函数，所以也应该视为一个随机变量。这里我们提示读者，没有任何理由可以把 y_t不看作一个现实中的金融时间序列变量，如股票收益率、利率等。但是，从回归方程（1.6）来观察 y_t，如果进一步假设方程中的各个系数给定或已知，我们可以把这样一个时序变量看作一种数据生成过程（data generating process，简记做 DGP）的产物。所谓数据生成过程，就是对被解释变量统计特性的一种完全刻画，本书第 3 章第 1 节将会进一步涉及并阐释金融计量学中经常用到的数据生产过程的概念。

注意，与随机变量紧密相关但又有区别的一个概念就是随机过程。当我们希望对一个金融时间序列进行分析时，通常把$\{y_t\}_1^T = (y_1, y_2, \cdots, y_t, y_{t+1} \cdots, y_T)$看作是一个随机过程的实现（realization）。宽泛地说，随机过程就是定义在一定概率空间上的一组具有相同特性的随机变量。为了方便说明，在金融计量分析中，通常可以使用 y_t既表示随机变量也代表相应的随机过程。为了不造成理解和说明上的混淆，本书中将采用这种传统。对于随机过程更严格的定义在本书第 3 章第 1 节有详细阐述。

1.2.3 随机分布

在金融计量分析中，除了常用的随机变量和随机过程的概念之外，还经常会涉及到随机分布的概念。要介绍这个概念，我们首先假定读者已经具有一定的统计学基础知识。然后，假设以向量形式组织的随机变量 X 和 Y，用$\Pr(X \subset \mathbb{R}^m, Y \subset \mathbb{R}^n)$表示 X 和 Y 在欧几里德空间$\mathbb{R}^m$和$\mathbb{R}^n$上对应的概率。再定义$x \in \mathbb{R}^m$是$\mathbb{R}^m$上的一点，同理定义 y。

这样，X 和 Y 的联合分布（joint distribution）就定义为：

$$F_{X,Y}(x, y; \vartheta) = \Pr(X \le x, Y \le y) \tag{1.7}$$

其中：ϑ为联合分布函数中的参数。假定X 与Y的联合概率密度函数(joint probability density function）为$f_{x,y}(x,y;\vartheta)$，并且严格有定义，则有：

$$F_{X,Y}(x,y;\vartheta)=\int_{-\infty}^{x}\int_{-\infty}^{y}f_{x,y}(w,z;\vartheta)dzdw \tag{1.8}$$

与联合分布相对的概念是边际分布（marginal distribution)。例如，X 的边际分布可以通过将联合分布中与 x 不相关的随机变量设为$+\infty$来获得，即：

$$F_X(x;\vartheta)=F_{X,Y}(x,+\infty,\cdots,+\infty;\vartheta) \tag{1.9}$$

当 X 是一个一维的随机变量而不是向量形式时，边际分布的定义就成为下面常见的形式：

$$F_X(x;\vartheta)=\Pr(X\le x;\vartheta) \tag{1.10}$$

公式（1.10）在统计学中也称为 X 的累积分布函数（Cumulative Distribution Function，简记做 CDF)，其取值范围在 0 与 1 之间。虽然 CDF 的概念稍微有些抽象，但是其在金融计量学中有着广泛的应用，特别是在计算统计量的 p-值过程中非常有用。例如，利用 F 分布的累积分布函数可以计算 F 检验统计量的 p-值。

谈及随机分布，另外一个相关的概念是条件分布。所谓条件分布，顾名思义，就是随机变量在给定条件下的分布。例如，给定$Y\le y$的条件，X 的条件分布可以定义为：

$$F_{X|Y\le y}(x;\vartheta)=\frac{\Pr(X\le x,Y\le y)}{\Pr(Y\le y)} \tag{1.11}$$

如果利用前面提到的概率密度函数的概念，还可以写成：

$$F_{x|y}(x;\vartheta)=\frac{F_{x,y}(x,y;\vartheta)}{F_y(y;\vartheta)} \tag{1.12}$$

其中：$F_y(y;\vartheta)$表示边际分布函数，并且满足：

$$F_y(y;\vartheta)=\int_{-\infty}^{+\infty}f_{x,y}(x,y;\vartheta)dx \tag{1.13}$$

这里提出两点需注意的问题。第一是公式（1.12）所刻画的关系经常被用在金融计量中的最大似然估计方法（maximum likelihood estimator，简记做 MLE）中。该方法的一个明显特征就是需要对涉及的随机变量（如回归方程中的随机扰动项）的统计分布有严格的假设，如正态分布假设等。有关 MLE 的知识将在本书的后续章节中做更多的介绍。第二个需要注意的是随机变量 X 与 Y 彼此之间是统计独立的，当且仅当条件分布与边际分布满足以下关系：

$$F_{x|y}(x;\vartheta)=F_x(x;\vartheta) \tag{1.14}$$

1.2.4　随机变量的期望与矩

随机变量的期望和矩是金融计量学习中非常有用的概念，而且这两个概念彼此依存度很高。注意，矩的英文对应名称为“moment”，可能与该英文单词常见的释义有些差别，这里特别提出以方便读者在阅读外版教材时不致产生混淆。

从统计学角度来说，一个随机变量 X 的 n 阶矩可以定义为：

$$E\left[(X-\mu)^n\right]=\int_{-\infty}^{+\infty}(x-\mu)^n f(x)dx \tag{1.15}$$

其中：μ表示 X 的均值，$E(\cdot)$表示期望操作符号。

这样，如果假设 X 的各阶矩都有定义，那么 n 取不同的值就对应着随机变量 X 的各阶矩。相应地，1 阶矩被称为均值或期望值，经常使用μ来表示，它度量了随机分布的中央位置。另外，随机变量的 2 阶矩被称为方差，经常用σ^2来表示，它衡量了 X 的变化幅度。方差的平方根称为标准差，用σ表示。均值和方差被广泛地用来刻画各种随机分布尤其是

正态分布的特性。

在金融计量学中，随着对非正态分布的研究的不断增多，随机变量的大于 2 阶的高阶矩也越来越多地引起研究者的兴趣。例如，3 阶矩又称为偏度（skewness），它度量了随机变量分布的非对称程度。如果偏度统计量为正值，说明 X 的分布具有右侧长尾特征，而负的偏度表明左侧长尾特征。另外，4 阶矩又称为峰度，其衡量随机变量分布的尖峰程度或平坦程度。作为经验值，对于一个服从正态分布的随机变量，其对应的偏度为 0，而峰度为 3。所以，可以通过偏度和峰度统计量来检验随机变量是否服从正态分布。著名的 Jarque-Bera 正态分布检验（Jarque and Bera, 1960）便是利用了这一思路设计的一种计量检验。Jarque-Bera 检验统计量定义为：

$$\text{Jarque}-\text{Bera}=\frac{N-k}{6}[S^2+\frac{(K-3)^2}{4}] \tag{1.16}$$

其中：N 代表样本大小，S 表示偏度，K 表示峰度，k 代表解释变量个数，原假设是待检验随机变量为正态分布。

对于正态分布的检验原理本书不再做过多的解释，这里总结一下随机变量的期望与矩之间的关系。考虑随机变量 $\varepsilon_t \sim iid(0,\sigma^2)$，其前 4 阶矩可以写出：

$$\begin{aligned}&E[\varepsilon_t]=0\\&E\left[(\varepsilon_t-E[\varepsilon_t])^2\right]=\text{var}[\varepsilon_t]=E[\varepsilon_t^2]=\sigma^2\\&E\left[(\varepsilon_t-E[\varepsilon_t])^3\right]=E[\varepsilon_t^3]\\&E\left[(\varepsilon_t-E[\varepsilon_t])^4\right]=E[\varepsilon_t^4]\end{aligned} \tag{1.17}$$

当然，如果进一步假设 ε_t 服从正态分布，那么第 3 和第 4 阶矩分别写作：

$$\begin{aligned}&E\left[(\varepsilon_t-E[\varepsilon_t])^3\right]=E[\varepsilon_t^3]=0\\&E\left[(\varepsilon_t-E[\varepsilon_t])^4\right]=E[\varepsilon_t^4]=3\end{aligned} \tag{1.18}$$

在金融时间序列分析中，前两阶矩是最为常用的。虽然上面介绍的矩的定义非常重要，但是在实际应用中，我们通常很难获得随机变量的“期望”，所以，随机变量的矩经常使用样本数据进行估计。例如，假定 T 代表随机变量 ε_t 的样本观测值的个数，那么样本均值可以写成：

$$\hat{\mu}=\frac{1}{T}\sum_{t=1}^{T}x_t \tag{1.19}$$

样本方差则为：

$$\hat{\sigma}^2=\frac{1}{T-1}\sum_{t=1}^{T}(x_t-\hat{\mu})^2 \tag{1.20}$$

作为一种运算符号，期望和方差运算操作符还具有以下性质(这里 A 表示常数，u_t 和 ε_t 表示均值为 0 的随机变量)：

$$\begin{aligned}&E(A\varepsilon_t)=AE[\varepsilon_t]\\&E[A+\varepsilon_t]=A+E[\varepsilon_t]\\&E[u_t+\varepsilon_t]=E[u_t]+E[\varepsilon_t]\\&\text{var}[A\varepsilon_t]=A^2\,\text{var}[\varepsilon_t]\\&\text{var}[A+\varepsilon_t]=\text{var}[\varepsilon_t]\\&\text{var}[u_t+\varepsilon_t]=E[u_t^2+2u_t\varepsilon_t+\varepsilon_t^2]=E[u_t^2]+2E[u_t\varepsilon_t]+E[\varepsilon_t^2]\end{aligned} \tag{1.21}$$

注意，从模型（1.21）所给出的最后一个等式可以看出，如果没有特殊的假设，一般来说 $E[u_t\varepsilon_t]\neq E[u_t]E[\varepsilon_t]$。事实上，$E[u_t\varepsilon_t]$ 表明的是两个随机变量之间的相关关系，其实质是 u_t