



21世纪统计学系列精品教材



应用时间序列分析

黄红梅◎编著

清华大学出版社





21世纪统计学系列精品教材



应用时间序列分析

黄红梅◎编著

清华大学出版社
北京

内 容 简 介

本书主要介绍了时间序列分析的一些经典和常用分析方法,主要包括 ARMA 模型、ADF 单位根检验、残差自回归模型、ARIMA 模型、季节性模型、GARCH 类模型、VAR 模型、协整和误差修正理论。本书在介绍基本理论的同时,注重厘清在建模实践过程中容易困扰初学者的典型问题。例如在 ADF 单位根检验过程中如何选择合适的检验模型来进行检验,趋势平稳过程和差分平稳过程的区分以及分别适用于何种建模方法,VAR 模型识别过程中变量先后次序的重要性,协整和误差修正模型中何时应加入截距或时间趋势项等。书中针对各种模型,以应用实例的方式介绍了 EViews 软件中的建模实践过程和注意事项,以帮助初学者掌握时间序列实证分析方法。同时,对每个应用实例还列示了相应的 R 语言实现命令及其结果,以方便习惯使用 R 语言的初学者掌握基本的时间序列建模方法。

本书可作为经济、统计、管理或金融类本科生或研究生的入门教材,也可作为时间序列分析实际工作者的应用型参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

应用时间序列分析/黄红梅编著. —北京:清华大学出版社,2016

21世纪统计学系列精品教材

ISBN 978-7-302-42278-5

I. ①应… II. ①黄… III. ①时间序列分析-高等学校-教材 IV. ①O211.61

中国版本图书馆 CIP 数据核字(2015)第 283603 号

责任编辑:苏明芳

封面设计:刘超

版式设计:魏远

责任校对:赵丽杰

责任印制:刘海龙

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62788951-223

印 装 者:北京密云胶印厂

经 销:全国新华书店

开 本:185mm×230mm 印 张:18.5 字 数:389千字

版 次:2016年3月第1版 印 次:2016年3月第1次印刷

印 数:1~3000

定 价:38.00元

产品编号:056256-01

前 言

随着时间序列理论研究的不断深入，自然科学和社会科学各领域应用时间序列分析技术的实践研究也日益广泛和深入。特别是在经济、统计、管理和金融等学科领域，应用时间序列分析方法对研究对象进行建模分析，已经成为这些学科开展实证研究的主流方法。本书在介绍时间序列分析方法过程中的实例主要以经济类的应用为主，介绍不同性质的时间序列数据应如何进行分析 and 建模。

时间序列分析实践应用能力的提高，一定是建立在必要的理论知识积累基础上的。因此，本书中的内容都是以理论知识和应用实例相结合的方式进行介绍。为了适应数学基础不是特别扎实的学习者的需求，本书在理论知识部分，重要公式的推导过程都力求详细。即使本书已经尽量考虑了学习者在数学基础上的差异，但为了更好地学习和掌握本书中的内容，学习者需要掌握必要的概率统计、回归分析和线性代数基础知识。本书几乎在每章中都设置有应用实例，目的是帮助实践研究者熟悉实际问题的研究方法和研究步骤，以及具体研究过程中的一些细节性问题。本书中的所有应用实例都借助 EViews 7 软件给出了详细的建模分析操作过程示例。同时，为了方便习惯使用 R 语言的学习者，也在每章的最后给出了所有实例在 R 语言中实现的编程命令和相应的输出结果（本书中的 R 语言编程命令和输出结果是借助于 RStudio 软件系统输出的）。为了更好地理解本书中应用实例的操作过程，学习者需要具备 EViews 软件或 R 语言操作的基础知识。由于 EViews 软件涉及更多的操作步骤（而 R 语言只是通过编程输出结果），为了方便不具备 EViews 软件操作基础的学习者，本书在书后附录中简要介绍了 EViews 软件操作的基础知识。

本书是在笔者讲授“应用时间序列分析”课程的过程中逐渐成稿的，书中的内容注重应用性兼顾理论阐释，定位为针对初学者的应用型参考书。虽然是针对初学者，但由于时间序列的波动性建模和多变量建模已经非常普遍，因此本书在重点介绍经典的 ARMA 建模方法之余，对 GARCH 类模型、VAR 模型、协整和误差修正理论也都进行了较为细致的理论阐述和实例分析，从而希望拓宽初学者对时间序列分析方法认识的广度，同时也希望他们对这些方法的应用不只是在一知半解基础上的错误应用，而是在对各项重要的细节问题思路清晰基础上的正确应用。

本书主要介绍了时间序列分析的一些经典和常用分析方法，主要包括 ARMA 模型、ADF 单位根检验、残差自回归模型、ARIMA 模型、季节性模型、GARCH 类模型、VAR

模型、协整和误差修正理论。本书在介绍基本理论的同时，注重厘清在建模实践过程中容易困扰初学者的典型问题。例如，在 ADF 单位根检验过程中如何选择合适的检验模型来进行检验，趋势平稳过程和差分平稳过程的区分以及分别适用于何种建模方法，VAR 模型识别过程中变量先后次序的重要性，协整和误差修正模型中何时应加入截距或时间趋势项等。特别值得强调的是，初学者在应用协整理论进行实证分析时往往存在很多误解，或者在具体的分析过程中对某些重要的细节性问题的理解存在偏误，从而得出错误的分析结论。本书对协整理论应用过程中的很多重要问题都特别进行了详细的阐述，例如 EG 协整检验过程中模型确定性回归变量的选择问题，不同变量作为被解释变量时 EG 协整检验出现检验结果矛盾时的处理方法，Johansen 协整检验的检验模型选择问题等。

参与本书编写工作的人员还有马利敏、孙利荣、袁中扬、陈娟、董亚娟，黄红梅对全书进行审阅、修改和定稿，并对全书负责。

在本书出版之际，我要感谢浙江工商大学统计与数学学院的赵卫亚老师和吴鑑洪老师给予的建议和帮助，也要感谢统计与数学学院的各位领导和同仁的关心和帮助。同时也要感谢母校吉林大学商学院的导师石柱鲜教授和各位老师，在他们的指导下我对时间序列分析的理论 and 实践产生了浓厚的兴趣。还要感谢东北财经大学的高铁梅教授和南开大学的张晓峒教授，他们曾多次在相关理论知识方面帮助我答疑解惑。清华大学出版社的苏明芳编辑也为本书的出版付出了大量的时间和精力，她严谨的工作态度令人印象深刻，与她的交流和合作是令人愉快的。最后，我要感谢家人对我的关爱、付出、理解和宽容。

由于作者水平有限，书中难免有疏漏、不当甚至错误之处，恳请同行专家和读者批评指正。联系邮箱：economicdoc09@163.com。

黄红梅

目 录

第 1 章 时间序列基础知识	1
1.1 时间序列的基本概念	1
1.1.1 时间序列的数字特征	1
1.1.2 时间序列自协方差和自相关系数的性质	2
1.2 平稳性	3
1.2.1 平稳性的定义	3
1.2.2 平稳时间序列的应用特性	4
1.3 白噪声过程	5
1.4 线性差分方程	6
1.4.1 滞后算子	6
1.4.2 差分算子	8
1.4.3 求解 p 阶线性差分方程的特征根法	9
1.4.4 求解 1 阶线性差分方程的特征根法和迭代法	12
习题及参考答案	13
参考文献	16
第 2 章 平稳时间序列模型	17
2.1 ARMA 模型的形式	17
2.1.1 ARMA 模型的典型形式	17
2.1.2 ARMA 模型的滞后算子多项式表示形式	18
2.1.3 ARMA 模型的传递形式和逆转形式	19
2.1.4 格林函数	19
2.2 ARMA 模型的平稳性条件	22
2.2.1 AR 模型的平稳性条件	22
2.2.2 MA 模型的平稳性条件	33
2.2.3 ARMA 模型的平稳性条件	35
2.3 MA 模型的可逆性条件	36
2.3.1 $MA(q)$ 模型的可逆性条件	36
2.3.2 $MA(1)$ 模型的可逆性条件	39

2.3.3	MA(2)模型的可逆性条件	39
2.4	ARMA 过程的自相关函数和 Yule-Walker 方程	41
2.4.1	AR(p)过程的自相关函数及其拖尾性	41
2.4.2	MA(q)过程的自相关函数及其截尾性	43
2.4.3	ARMA(p, q)过程的自相关函数及其拖尾性	44
2.5	ARMA 过程的偏自相关函数	45
2.5.1	AR(p)过程的偏自相关函数及其截尾性	46
2.5.2	MA(q)过程的偏自相关函数及其拖尾性	48
2.5.3	ARMA(p, q)过程的偏自相关函数及其拖尾性	49
2.6	Box-Jenkins 建模方法	49
2.7	ARMA 模型的预测	53
2.7.1	最小均方误差预测	53
2.7.2	条件期望	54
2.7.3	预测误差	54
2.7.4	AR(p)过程的预测	55
2.7.5	MA(q)过程的预测	56
2.7.6	ARMA(p, q)过程的预测	57
2.8	实例应用	60
2.9	补充内容	66
	习题及参考答案	78
	参考文献	81
第 3 章	单位根检验	82
3.1	典型的平稳和非平稳过程	82
3.1.1	零均值平稳过程和随机游走过程	82
3.1.2	非零均值平稳过程和带漂移的随机游走过程	84
3.1.3	趋势平稳过程和趋势非平稳过程	85
3.2	趋势平稳和差分平稳	86
3.2.1	趋势平稳过程	86
3.2.2	差分平稳过程	87
3.2.3	过度差分	88
3.3	单位根检验	88
3.3.1	DF 检验	89

3.3.2	ADF 检验	91
3.3.3	DF 和 ADF 检验中几个值得注意的问题	92
3.3.4	其他类型的单位根检验	97
3.4	实例应用	97
3.5	补充内容	104
3.5.1	例 3-1 的相关 R 语言命令及结果	104
3.5.2	例 3-2 的相关 R 语言命令及结果	110
	习题及参考答案	118
	参考文献	119
第 4 章	趋势和季节性建模	121
4.1	确定性趋势建模	121
4.2	随机趋势建模	127
4.3	季节性建模	127
4.4	补充内容	134
4.4.1	例 4-1 的相关 R 语言命令及结果	134
4.4.2	例 4-2 的相关 R 语言命令及结果	141
	习题及参考答案	149
	参考文献	150
第 5 章	条件异方差模型	151
5.1	异方差的定义	151
5.2	ARCH 过程	152
5.3	GARCH 过程	154
5.3.1	GARCH(p, q)模型	154
5.3.2	条件方差的预测	155
5.4	GARCH 过程的扩展模型	156
5.4.1	GARCH-M 模型	156
5.4.2	IGARCH 模型	157
5.4.3	EGARCH 模型	157
5.4.4	TGARCH 模型	158
5.4.5	GARCH 类过程特点比较	159
5.5	条件异方差性的诊断检验	160

5.6 实例应用	161
5.7 补充内容	168
习题及参考答案	181
参考文献	187
第 6 章 向量自回归模型	188
6.1 VAR 模型的标准形式	188
6.2 VAR 模型的平稳性条件	190
6.3 VAR 模型的估计和识别	192
6.3.1 VAR 模型的估计	192
6.3.2 VAR 模型的识别——2 维 1 阶情况	193
6.3.3 VAR 模型的识别—— n 维 1 阶情况	195
6.4 脉冲响应函数	196
6.4.1 Cholesky 分解下的脉冲响应	196
6.4.2 相关系数对 Cholesky 分解中变量次序重要性的影响	198
6.5 方差分解	203
6.6 Granger 因果关系检验	204
6.7 实例应用	205
6.8 补充内容	212
习题及参考答案	221
参考文献	224
第 7 章 协整和误差修正	225
7.1 协整理论	225
7.2 Engle-Granger 协整检验	227
7.2.1 EG 协整检验原理	227
7.2.2 EG 协整检验实例	229
7.3 Johansen 协整检验	234
7.3.1 Johansen 协整检验原理	234
7.3.2 向量误差修正模型	238
7.3.3 VECM 中的截距项	241
7.3.4 模型滞后阶数的选择	245
7.3.5 Johansen 协整检验统计量	246

7.3.6 Johansen 协整检验实例.....	247
7.4 伪回归.....	251
7.5 补充内容.....	252
习题及参考答案.....	262
参考文献.....	264
附录 A EViews 软件快速入门.....	266
A.1 主界面窗口.....	266
A.2 工作文件.....	267
A.3 常用对象.....	269
A.3.1 序列.....	270
A.3.2 方程.....	271
A.3.3 组.....	273
A.3.4 向量自回归.....	274
附录 B 实例数据.....	276
附录 C R 语言函数索引.....	285

第 1 章 时间序列基础知识

本章导读

时间序列分析的一个重要特点，是对平稳时间序列和非平稳时间序列分别采取不同的建模策略。为了更好地理解平稳时间序列模型的性质，首先需要对平稳性有一个充分的认识；而了解线性差分方程解的表示，对于深入理解平稳性及模型相关性质都具有重要意义。

本章结构如下：1.1 节介绍时间序列的基本概念；1.2 节介绍平稳性的含义；1.3 节介绍在时间序列建模中起到基石作用的白噪声过程；1.4 节简要介绍线性差分方程及其求解。

1.1 时间序列的基本概念

时间序列 (time series) 是按照时间顺序排列的一组随机变量。时间序列与随机过程紧密相关，在时间序列的理论研究过程中经常将其理解为一个随机过程。随机过程 (stochastic process) 是一组有序的随机变量，可以记为 $\{y(t), t \in T\}$ 。随机过程一般是定义在连续集合上的，而定义在离散集合上的随机过程则通常称为时间序列。离散的时间集合 T 可以表示为 $T = \{\dots, -2, -1, 0, 1, 2, \dots\}$ ，此时 $y(t)$ 是离散时间 t 的随机函数，时间序列通常表示为 $\{y_t, t = \dots, -2, -1, 0, 1, 2, \dots\}$ 。

时间序列在特定时间段上的观测样本可以视为随机过程的一次实现，通常称为样本序列，记为 $\{y_0, y_1, y_2, \dots, y_T\}$ 。理论上说，时间序列可以有无限个观测时间点，然而从实际可获得的样本数据来看，样本序列都是有限的。更加重要的是，由于时间的不可重复性，时间序列通常仅有一次实现，即只有一个样本序列。因此时间序列的经验研究的一个显著特点是，只能在唯一可观测到的样本序列的基础上来推测时间序列的总体特性。

简便起见，本书中有些时候将随机过程简称为过程，将时间序列简称为序列，都简记为 $\{y_t\}$ 。在不会导致混淆的情况下，样本序列也用 $\{y_t\}$ 表示。

1.1.1 时间序列的数字特征

我们用 y_t 表示时间序列 $\{y_t\}$ 中 t 时刻的随机变量，用 y_k 表示 k 时刻的随机变量，并作如下定义：

$$\mu_t = E(y_t) \quad (1-1)$$

$$\sigma_t^2 = E(y_t - \mu_t)^2 \quad (1-2)$$

$$\gamma(t, k) = \text{Cov}(y_t, y_k) = E(y_t - \mu_t)(y_k - \mu_k) \quad (1-3)$$

$$\rho(t, k) = \frac{\gamma(t, k)}{\sqrt{\sigma_t^2 \times \sigma_k^2}} = \frac{\gamma(t, k)}{\sigma_t \times \sigma_k} \quad (1-4)$$

式(1-1)和式(1-2)表明 y_t 的均值定义为 μ_t 、方差定义为 σ_t^2 。这里用下标 t 来标识特定时间点上随机变量的均值和方差。当 t 取尽所有可能时间点时,便形成了对应于时间序列 $\{y_t\}$ 的均值序列和方差序列。

式(1-3)表明 y_t 和 y_k 之间的自协方差定义为 $\gamma(t, k)$, 这里之所以称其为“自”是由于它是取自于同一个时间序列不同时间点随机变量间的协方差。当 t 和 k 取尽所有可能时间点时,形成关于 t 和 k 的二元离散函数 $\gamma(t, k)$, 称为自协方差函数。应当看到,当 $k = t$ 时,自协方差就等于 t 时刻的方差,即

$$\gamma(t, t) = \sigma_t^2 \quad (1-5)$$

式(1-4)表明 y_t 和 y_k 之间的自相关系数定义为 $\rho(t, k)$, 该定义与随机变量间相关系数的定义类似,都揭示了标准化的相关关系。

1.1.2 时间序列自协方差和自相关系数的性质

1. 时间序列自协方差的性质

(1) 对称性

$$\gamma(t, k) = \gamma(k, t) \quad (1-6)$$

(2) 非负定性

对任意 m 个时间点 t_1, t_2, \dots, t_m , 自协方差矩阵 Γ_m 是对称非负定矩阵。

$$\Gamma_m = \begin{bmatrix} \gamma(t_1, t_1) & \gamma(t_1, t_2) & \cdots & \gamma(t_1, t_m) \\ \gamma(t_2, t_1) & \gamma(t_2, t_2) & \cdots & \gamma(t_2, t_m) \\ \vdots & \vdots & \vdots & \vdots \\ \gamma(t_m, t_1) & \gamma(t_m, t_2) & \cdots & \gamma(t_m, t_m) \end{bmatrix} \quad (1-7)$$

2. 时间序列自相关系数的性质

(1) 规范性

$$\rho(t, t) = 1 \text{ 且 } |\rho(t, k)| \leq 1 \quad (1-8)$$

(2) 对称性

$$\rho(t, k) = \rho(k, t) \quad (1-9)$$

(3) 非负定性

对任意 m 个时间点 t_1, t_2, \dots, t_m , 自相关系数矩阵 \mathbf{P}_m 是对称非负定矩阵。

$$\mathbf{P}_m = \begin{bmatrix} \rho(t_1, t_1) & \rho(t_1, t_2) & \cdots & \rho(t_1, t_m) \\ \rho(t_2, t_1) & \rho(t_2, t_2) & \cdots & \rho(t_2, t_m) \\ \vdots & \vdots & \vdots & \vdots \\ \rho(t_m, t_1) & \rho(t_m, t_2) & \cdots & \rho(t_m, t_m) \end{bmatrix} \quad (1-10)$$

1.2 平稳性

在进行时间序列分析时, 针对时间序列的平稳 (stationary) 和非平稳特性需要采取不同的建模方法进行研究, 因此区分研究对象是平稳时间序列还是非平稳时间序列是时间序列分析的首要步骤。

时间序列分析理论中有两种平稳性定义, 即所谓严平稳 (strictly stationary) 和弱平稳 (weakly stationary)。严平稳也称强平稳 (strongly stationary), 它是在时间序列中各时刻随机变量联合分布函数的基础上定义的更为严格的平稳性。粗略地说, 严平稳时间序列的所有统计性质都不随时间的变化而改变。

弱平稳也称协方差平稳 (covariance stationary)、二阶平稳 (second-order stationary) 或宽平稳 (wide-sense stationary), 它是在时间序列二阶矩基础上定义的平稳性。简单来说, 弱平稳时间序列的一阶矩和二阶矩不随时间的变化而改变。

一般来说, 满足严平稳的序列也具有弱平稳性, 但严平稳却不能全部涵盖弱平稳。例如, 如果一个严平稳时间序列不存在二阶矩或一阶矩 (如柯西分布), 则它就不满足弱平稳性。

另外, 一个弱平稳序列的分布不一定满足不随时间的变化而改变, 因此弱平稳序列并不一定严平稳。但如果时间序列为正态序列, 则由于二阶矩描述了正态分布的所有统计性质, 因此弱平稳的正态序列也是严平稳的。

1.2.1 平稳性的定义

【定义 1.1】 如果时间序列 $\{y_t\}$ 的二阶矩有限, 且满足

$$E(y_t) = E(y_{t-j}) = \mu \quad (1-11)$$

$$\text{Var}(y_t) = \text{Var}(y_{t-j}) = \sigma^2 \quad (1-12)$$

$$\text{Cov}(y_t, y_{t-s}) = \text{Cov}(y_{t-j}, y_{t-j-s}) = \gamma_s \quad (1-13)$$

对所有的 t, j 和 s 成立, 其中 μ, σ^2 和 γ_s 均为常数, 则称时间序列 $\{y_t\}$ 是弱平稳的。

弱平稳定义中, 式 (1-11) 和式 (1-12) 表明弱平稳时间序列具有有限的常数均值和方差, 式 (1-13) 表明弱平稳时间序列的自协方差只与时滞 s 有关, 而与时间的起始位置 t 无关。概括来说, 弱平稳时间序列的一阶矩和二阶矩都是不随时间的变化而改变的常数。

由于弱平稳时间序列的自协方差只与时滞 s 有关, 而与 t 无关, 因此式 (1-13) 中将二元函数形式的自协方差 $\text{Cov}(y_t, y_{t-s})$ 简记为仅与时滞 s 相关的一元函数形式 γ_s 。例如 γ_1 表示滞后 1 期的自协方差, γ_2 表示滞后 2 期的自协方差等。值得注意的是, 当 $s=0$ 时, γ_0 就等同于方差, 有

$$\gamma_0 = \sigma^2 \quad (1-14)$$

此外, 与自协方差类似, 平稳时间序列的自相关系数也只与时滞 s 有关, 因此平稳时间序列的自相关系数也可以简记为与时滞 s 相关的一元函数形式 ρ_s , 并有

$$\rho_s = \frac{\gamma_s}{\gamma_0} = \frac{\gamma_s}{\sigma^2} \quad (1-15)$$

由于平稳时间序列的自相关系数是时滞 s 的函数, 因此通常也称 ρ_s 为自相关函数 (Auto-Correlation Function, ACF), ρ_s 对时滞 s 作图通常称为自相关图 (correlogram)。

在本书所涉及的时间序列建模理论中, 只考虑弱平稳性即可, 因此本书中后续涉及的平稳性都指弱平稳性。

1.2.2 平稳时间序列的应用特性

从平稳时间序列的定义可以看到, 平稳时间序列 $\{y_t\}$ 中任一时间点上随机变量的一阶矩和二阶矩都为相同的有限常数。虽然如此, 这些常数一阶矩和二阶矩的估计却仍是一个问题。因为通常情况下, 由于时间的不可重复性, 一个时间序列仅能够获得一个观测值序列, 即时间序列中每个时间点上的随机变量仅有一个单一样本。通常情况下, 通过每个时间点上所能观测到的单一的样本来估计对应随机变量的一阶矩和二阶矩是不可行的。

幸运的是, 平稳时间序列在一定条件下^①, 可以通过所有时间点上各样本的集合来估计这些常数一阶矩和二阶矩。例如, 假设平稳时间序列 $\{y_t\}$ 有 T 个观测值, 分别用

^① 协方差平稳的时间序列在满足遍历性 (ergodicity) 要求时, 其观测值序列各时间点上单一样本集合的特定函数形式的时间平均, 才依概率收敛于对应的总体数字特征。例如, 当 $T \rightarrow \infty$ 时, $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ 依概率收敛于 $E(y_t) = \mu$, 则称该协方差平稳序列是关于均值遍历的; 当 $T \rightarrow \infty$ 时, $\frac{1}{T-s} \sum_{t=1}^{T-s} (y_t - \bar{y})(y_{t+s} - \bar{y})$ 依概率收敛于 $\text{Cov}(y_t, y_{t+s}) = \gamma_s$, 则称该协方差平稳序列是关于二阶矩遍历的。协方差平稳时间序列关于均值遍历和关于二阶矩遍历的条件参见汉密尔顿^[1]所著的《时间序列分析》。

y_1, y_2, \dots, y_T 表示, 用 $\hat{\mu}$, $\hat{\sigma}^2$, $\hat{\gamma}_s$ 和 $\hat{\rho}_s$ 分别表示 μ , σ^2 , γ_s 和 ρ_s 的估计, 则有^①

$$\hat{\mu} = \bar{y} = \frac{1}{T} \sum_{t=1}^T y_t \quad (1-16)$$

$$\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^T (y_t - \bar{y})^2 \quad (1-17)$$

$$\hat{\gamma}_s = \hat{\gamma}_{-s} = \frac{1}{T-s} \sum_{t=1}^{T-s} (y_t - \bar{y})(y_{t+s} - \bar{y}) \quad (1-18)$$

$$\hat{\rho}_s = \hat{\rho}_{-s} = \frac{\hat{\gamma}_s}{\hat{\gamma}_0} \quad (1-19)$$

其中, $\hat{\mu}$, $\hat{\sigma}^2$, $\hat{\gamma}_s$ 和 $\hat{\rho}_s$ 分别为样本均值、样本方差、样本自协方差和样本自相关函数, 式 (1-18) 和式 (1-19) 还表明样本自协方差和样本自相关函数具有对称性。

1.3 白噪声过程

时间序列分析的主要内容是对时间序列建模, 而时间序列模型中的随机性往往是通过白噪声 (white noise) 过程^②来引入的。白噪声过程可谓是时间序列模型构建的基石。

【定义 1.2】若 $\{\varepsilon_t\}$ 满足零均值、同方差和非自相关, 即

$$E(\varepsilon_t) = 0 \quad (1-20)$$

$$\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 \quad (1-21)$$

$$E(\varepsilon_t \varepsilon_{t-s}) = 0 \quad (1-22)$$

对所有的 t 和 $s \neq t$ 成立, 则称 $\{\varepsilon_t\}$ 为白噪声过程, 通常可记作 $\varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2)$ 。特别地, 如果 ε_t 服从正态分布, 则称 $\{\varepsilon_t\}$ 为正态白噪声过程或高斯白噪声过程。

从白噪声过程的定义可以看出, 它满足平稳性条件, 因此白噪声过程是一种特殊的平稳过程。白噪声过程的显著特点是高度的随机性, 也称纯随机性, 即各时刻随机变量之间互不相关, 因此就没有必要构建时间序列模型再去研究其相关关系。一般情况下, 对一个时间序列进行研究, 当各时刻随机变量之间所有的相关关系都已经通过建模被识别后, 剩下的无须进一步研究的部分通常都是白噪声过程, 也就是在这种意义上, 白噪声过程成了时间序列模型的基本构件。

^① 大样本情形下, 自协方差的估计也可为 $\hat{\gamma}_s = \frac{1}{T} \sum_{t=1}^{T-s} (y_t - \bar{y})(y_{t+s} - \bar{y})$ 。

^② 之所以称之为白噪声过程, 是由于白噪声过程的功率谱密度在整个频域内均匀分布, 即在所有频率上的能量相同, 这与“白光”的特性类似, 因此而得名。

本书中接下来都用 $\{\varepsilon_t\}$ 来表示白噪声过程,如果涉及不同的白噪声过程,则用 $\{\varepsilon_{1t}\}$ 和 $\{\varepsilon_{2t}\}$ 等来表示。

【例 1-1】EViews 软件中通过在命令行输入命令:

```
series e=nrnd
```

可以生成一个服从 $N(0,1)$ 的标准正态白噪声过程的样本序列 e , 该序列的时序图^①如图 1-1 所示。

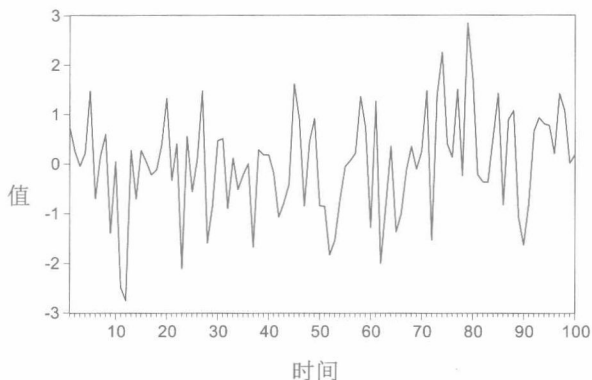


图 1-1 标准正态白噪声过程样本序列时序图

1.4 线性差分方程

时间序列建模的显著特点是关注相继的一些随机变量之间的相关关系,也因为如此,时间序列模型一般都为动态模型。这一建模特点使得线性时间序列模型在很大程度上与线性差分方程类似,事实上可以将多数线性时间序列模型看作包含随机成分的线性差分方程。因此在介绍时间序列建模理论之前,首先需要介绍线性差分方程及其求解,这对于理解时间序列模型的各项统计性质及各种分析方法都具有非常重要的意义。

1.4.1 滞后算子

假设已知时间序列 $\{y_t\}$ 和 $\{z_t\}$ 有如下关系

$$z_t = y_{t-1} \quad (1-23)$$

^① 时序图是指以时间点为横轴标识,对应时间序列为纵轴标识的顺序曲线图。

这一关系式也可以用滞后算子（也称延迟算子） L 表示为

$$z_t = Ly_t \quad (1-24)$$

式(1-23)和式(1-24)意味着在各时间点上, 时间序列 $\{y_t\}$ 和 $\{z_t\}$ 的对应关系如表1-1所示。

表 1-1 $z_t = Ly_t$ 关系下时间序列 $\{y_t\}$ 和 $\{z_t\}$ 的对应关系

时期 序列	1	2	3	4	5	6	7	8	9	10	...
y_t	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	...
z_t	—	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	...

因此, 滞后算子 L 的作用就是将时间序列逐项推后一期。关于滞后算子有下面的一些关系式成立:

$$Ly_t = y_{t-1} \quad (1-25)$$

$$L^j y_t = y_{t-j} \quad (1-26)$$

其中, j 为整数。则根据式(1-26), 当 $j=0$ 时, $L^0 y_t = y_t$; 当 $j=-i$ 时, $L^{-i} y_t = y_{t+i}$ 。

另外, 滞后算子具有如下性质。

(1) 对常数施加滞后算子仍为常数, 即

$$Lc = c \quad (1-27)$$

其中, c 表示常数。

(2) 滞后算子适用分配率, 即

$$(L^i + L^j)y_t = L^i y_t + L^j y_t = y_{t-i} + y_{t-j} \quad (1-28)$$

(3) 滞后算子适用结合率, 即

$$L^i L^j y_t = L^i (L^j y_t) = L^i y_{t-j} = y_{t-i-j} \quad (1-29)$$

滞后算子还可以通过线性运算, 构造滞后算子多项式来对时间序列进行更加复杂的运算。典型的 p 阶滞后算子多项式有如式(1-30)等号右边的形式, 并可将其简记为 $A(L)$ 。

$$A(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_p L^p \quad (1-30)$$

则滞后算子多项式 $A(L)$ 施加于时间序列 $\{y_t\}$ 时有

$$\begin{aligned} A(L)y_t &= (\alpha_0 + \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_p L^p)y_t \\ &= \alpha_0 y_t + \alpha_1 Ly_t + \alpha_2 L^2 y_t + \cdots + \alpha_p L^p y_t \\ &= \alpha_0 y_t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} \end{aligned} \quad (1-31)$$