



“十二五”国家重点图书出版规划项目
电子与信息工程系列

TIME SERIES ANALYSIS AND MODERN SPECTRUM ESTIMATION

时间序列分析与现代谱估计

● 冀振元 编著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS



“十二五”国家重点图书出版规划项目
电子与信息工程系列

TIME SERIES ANALYSIS AND MODERN SPECTRUM ESTIMATION

时间序列分析与现代谱估计

● 冀振元 编著



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内容简介

本书系统地讲述了时间序列分析的基本理论、建模步骤、预测方法以及现代谱估计的特点和相关知识。全书共分6章。第1章绪论,介绍时间序列分析的重要性、时间序列分析的发展及应用等内容;第2章介绍时间序列模型建立前的动态数据预处理,包括平稳性检验、正态性检验、独立性检验、周期性检验、趋势项检验等内容;第3章介绍常用的时间序列模型,包括自回归(AR)模型、移动平均(MA)模型、自回归移动平均(ARMA)模型、ARMA模型的特性、平稳时间序列模型的建立、平稳时间序列预测等内容;第4章介绍经典谱分析的基本方法,包括自相关函数的估计、经典谱估计的直接法、间接法及改进方法等;第5章介绍现代谱估计中的常用方法,包括线性预测法、Burg法、Prony法、多信号分类(MUSIC)法、基于旋转不变技术的信号参数估计(ESPRIT)法、最小范数法等;第6章介绍时间序列分析与谱估计常用的软件及实验相关内容。

本书可作为通信、电子信息、自动控制、概率统计等相关专业的研究生教材,也可作为相关技术人员在时间序列分析与谱估计方面研究的理论基础参考书。

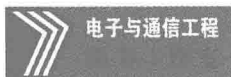
图书在版编目(CIP)数据

时间序列分析与现代谱估计/冀振元编著. —哈尔滨:哈尔滨工业大学出版社,2016.3

ISBN 978-7-5603-5794-2

I. ①时… II. ①冀… III. ①时间序列分析②谱估计
IV. ①O211.61②TN911.23

中国版本图书馆 CIP 数据核字(2016)第 003954 号



责任编辑 许雅莹

封面设计 刘洪涛

出版发行 哈尔滨工业大学出版社

社 址 哈尔滨市南岗区复华四道街10号 邮编 150006

传 真 0451-86414749

网 址 <http://hitpress.hit.edu.cn>

印 刷 黑龙江省地质测绘印制中心印刷厂

开 本 787mm×1092mm 1/16 印张 12.25 字数 260千字

版 次 2016年3月第1版 2016年3月第1次印刷

书 号 ISBN 978-7-5603-5794-2

定 价 34.00元

(如因印装质量问题影响阅读,我社负责调换)

“十二五”国家重点图书 电子与信息工程系列

编 审 委 员 会

顾 问 张乃通

主 任 顾学迈

副 主 任 张 晔

秘 书 长 赵雅琴

编 委 (按姓氏笔画排序)

王 钢 邓维波 任广辉 沙学军

张钧萍 吴芝路 吴 群 谷延锋

孟维晓 赵洪林 赵雅琴 姜义成

郭 庆 宿富林 谢俊好 冀振元

序

FOREWORD

教材建设一直是高校教学建设和教学改革的主要内容之一。针对目前高校电子与信息工程教材存在的基础课教材偏重数学理论,而数学模型和物理模型脱节,专业课教材对最新知识增长点和研究成果跟踪较少等问题,及创新型人才的培养目标和各学科、专业课程建设全面需求,哈尔滨工业大学出版社与哈尔滨工业大学电子与信息工程学院的各位老师策划出版了电子与信息工程系列精品教材。

该系列教材是以“寓军于民,军民并举”为需求前提,以信息与通信工程学科发展为背景,以电子线路和信号处理知识为平台,以培养基础理论扎实、实践动手能力强的创新型人才为主线,将基础理论、电信技术实际发展趋势、相关科研开发的实际经验密切结合,注重理论联系实际,将学科前沿技术渗透其中,反映电子信息领域最新知识增长点和研究成果,因材施教,重点加强学生的理论基础水平及分析问题、解决问题的能力。

本系列教材具有以下特色:

(1)强调平台化完整的知识体系。该系列教材涵盖电子与信息工程专业技术理论基础课程,对现有课程及教学体系不断优化,形成以电子线路、信号处理、电波传播为平台课程,与专业应用课程的四个知识脉络有机结合,构成了一个通识教育和专业教育的完整教学课程体系。

(2)物理模型和数学模型有机结合。该系列教材侧重在经典理论与技术的基础上,将实际工程实践中的物理系统模型和算法理论模型紧密结合,加强物理概念和物理模型的建立、分析、应用,在此基础上总结牵引出相应的数学模型,以加强学生对算法理论的理解,提高实践应用能力。

(3)宽口径培养需求与专业特色兼备。结合多年来有关科研项目的科研经验及丰硕成果,以及紧缺专业教学中的丰富经验,在专业课教材编写过程中,在兼顾电子与信息工程毕业生宽口径培养需求的基础上,突出军民兼用特色,在

满足一般重点院校相关专业理论技术需求的基础上,也满足军民并举特色的要求。

电子与信息工程系列教材是哈尔滨工业大学多年来从事教学科研工作的各位教授、专家们集体智慧的结晶,也是他们长期教学经验、工作成果的总结与展示。同时该系列教材的出版也得到了兄弟院校的支持,提出了许多建设性的意见。

我相信:这套教材的出版,对于推动电子与信息工程领域的教学改革、提高人才培养质量必将起到重要推动作用。

哈尔滨工业大学教授 张乃通
中国工程院院士



2010 年 11 月于哈工大

前言

PREFACE

时间序列分析是分析历史数据、建立模型、预测发展趋势最强有力的工具之一。它是利用随机过程理论和数理统计学的方法,研究随机数据序列所遵从的统计规律,以用于解决实际问题。

本书系统地讲述了时间序列分析的基本理论、建模步骤、预测方法以及现代谱估计的特点和相关知识。全书共分 6 章。

第 1 章绪论,介绍时间序列分析的重要性、时间序列分析的发展及应用等内容。

第 2 章介绍时间序列模型建立前的动态数据预处理,包括平稳性检验、正态性检验、独立性检验、周期性检验、趋势项检验等内容。

第 3 章内容包括:常用的时间序列模型:自回归(AR)模型、移动平均(MA)模型、自回归移动平均(ARMA)模型;ARMA 模型的特性:格林函数和平稳性、逆函数和可逆性、自协方差函数;平稳时间序列模型的建立:模型的识别、模型定阶、模型参数估计、模型的适应性检验;平稳时间序列预测:正交投影预测、条件期望预测、适时修正预测、指数平滑预测。

第 4 章介绍经典谱分析的基本方法,包括自相关函数的估计、经典谱估计的直接法、间接法及改进方法等。为后面各章节介绍现代谱估计的知识打下基础。

第 5 章介绍现代谱估计中的常用方法,包括线性预测法、Burg 法、Prony 法、多信号分类(MUSIC)法、基于旋转不变技术的信号参数估计(ESPRIT)法、最小范数法等。

第 6 章介绍时间序列分析与谱估计常用的软件及实验相关内容,以加强对学生实践环节的培养。

本书在编写过程中汲取了多本国内外优秀图书的精华,在这里向这些图书的作者表示感谢!

因作者水平有限,且编写时间仓促,书中难免有疏漏和不足之处,望广大读者谅解和批评指正!

编 者

2016 年 1 月

目录

CONTENTS

第 1 章 绪 论	1
1.1 时间序列分析的重要性	1
1.2 时间序列分析与随机过程理论的区别	1
1.3 时间序列分析方法的起源与发展	2
1.4 时间序列分析的应用领域	2
第 2 章 动态数据预处理	4
2.1 平稳性检验	4
2.2 正态性检验	8
2.3 独立性检验	13
2.4 周期性检验	14
2.5 趋势项检验	14
习 题	20
第 3 章 时间序列模型	21
3.1 一阶自回归(AR)模型	21
3.2 一般自回归模型	24
3.3 移动平均(MA)模型	25
3.4 自回归移动平均(ARMA)模型	26
3.5 ARMA 模型的特性	27
3.6 平稳时间序列模型的建立	52
3.7 平稳时间序列预测	75
习 题	88
第 4 章 经典谱分析	89
4.1 功率谱估计概述	89
4.2 自相关函数的估计	92
4.3 经典谱估计的基本方法	97
4.4 直接法和间接法估计的质量	100

4.5 直接法估计的改进	119
习 题	124
第 5 章 现代谱估计	125
5.1 引 言	125
5.2 自回归(AR)方法	125
5.3 输入数据处理	128
5.4 Burg 法	131
5.5 Prony 法	134
5.6 使用最小二乘途径的 Prony 法	137
5.7 特征向量和特征值	138
5.8 MUSIC 方法	140
5.9 ESPRIT 法	143
5.10 最小范数法	145
5.11 用离散傅里叶变换的最小范数法	148
习 题	149
第 6 章 时间序列分析与谱估计软件及实验指导	150
6.1 时间序列分析软件——EViews	150
6.2 MATLAB 介绍	154
6.3 时间序列分析及谱估计实验	181
参考文献	184

1.1 时间序列分析的重要性

人类为了探索周围的世界,常常依时间发展的先后顺序对某一事物进行观测。人们的一切活动,其根本目的无不在于认识和改造客观世界。时间序列分析(Time Series Analysis)不仅可以从数量上揭示某一现象的发展变化规律或从动态的角度刻画某一现象和其他现象之间的内在数量关系及其变化规律性,达到认识客观世界的目的,而且运用时间序列模型还可以预测和控制现象的未来行为,修正或重新设计系统以达到利用和改造客观世界之目的。

从统计学的内容来看,统计所研究和处理的是一批有“实际背景”的数据,尽管数据的背景和类型各不相同,但从数据的形成来看,无非是横剖面数据和纵剖面数据两类(或者称为静态数据和动态数据)。横剖面数据是由若干相关现象在某一时点上所处的状态组成的,它反映一定时间、地点等客观条件下诸相关现象之间存在的内在数值联系。研究这种数据结构的统计方法是多元统计分析。纵剖面数据是由某一现象或若干现象在不同时刻上的状态所形成的数据,它反映的是现象以及现象之间关系的发展变化规律性。研究这种数据的统计方法就是时间序列分析,这足以看出时间序列分析的重要性及其应用的广泛性。

1.2 时间序列分析与随机过程理论的区别

时间序列分析是概率统计学科的一个分支,它是运用概率统计的理论和方法来分析随机数据序列(或称动态数据序列),并对其建立数学模型,进行参数估计,对模型定阶,以及进一步应用于预报、预测、自适应控制、最佳滤波等诸多方面。

时间序列分析方法与随机过程理论有所区别,前者是先对实测数据建立数学模型,并在此基础上进一步分析随机数据的统计特性;后者是在对实测数据统计所得的先验概率知识基础上来分析其统计特性。由于人们所能获得的实测数据总是有限的,而理论上的先验概率要求在无限多的样本数据基础上统计才能获得,因此实际上我们能够获得的先验概率只



能是在一定置信度条件下的近似,亦即尽量接近真实的概率(密度)分布,这是随机过程理论和方法在实际应用时的困难。时间序列分析方法可以克服这一困难,它是在有限个样本数据总量的情况下建立起相当精确的数学模型,从而获得具有一定精度(用模型误差方差来表示)的统计特性,与真实结果非常接近,因此在实际应用时比较方便,可操作性较好。

人们将随机过程称为大样本理论,它是用多维概率分布来描述动态数据,因为多维概率分布是要建立在无限多样本数据的统计基础上,因此称为大样本理论。时间序列分析则可以从有限的样本数据中拟合成具有一定精度的时间序列模型,因此它又可称为小样本理论。

总之,随机过程分析方法在理论上严谨求实,但操作性较差;而时间序列分析方法在使用时方便实用,但是,要想建立精度相当高的时序模型不仅要求模型参数最佳地估计,而且模型阶数也要合适,因此建模过程也是相当复杂的。这两种对随机数据序列的分析方法都有各自的研究和应用领域,应视不同的分析对象和要求而定。

1.3 时间序列分析方法的起源与发展

时间序列分析起源于预测,尤其是对市场经济方面的预测,也就是说,时间序列分析本来的就是对某一事物或现象进行预测。时间序列分析方法最早起源于1927年,数学家耶尔(Yule)提出建立自回归(AR)模型来预测市场变化的规律。在1931年,数学家瓦尔格(Walker)在AR模型的启发下,建立了滑动平均(MA)模型和自回归、滑动平均(ARMA)混合模型,初步奠定了时间序列分析方法的基础,当时主要应用在经济分析和市场预测领域。20世纪60年代,时间序列分析理论和方法迈入了一个新的阶段,伯格(Burg)在分析地震信号时最早提出最大熵谱(MES)估计理论,后来有人证明AR模型的功率谱估计与最大熵谱估计是等效的,并称之为现代谱估计。它克服了用传统的傅里叶功率谱分析(又称经典谱分析)所带来的分辨率不高和频率漏泄严重等固有的缺点,从而使时间序列分析方法不仅在时间域内得到应用,而且扩展到频率域内,得到更加广泛的应用,特别是在各种工程领域内应用功率谱的概念更加方便和普通。到20世纪70年代以后,随着信号处理技术的发展,时间序列分析方法不仅在理论上更趋完善,尤其是在参数估计算法、定阶方法及建模过程等方面都得到了许多改进,进一步地迈向实用化,各种时间序列分析软件也不断涌现,逐渐成为分析随机数据序列不可缺少的有效工具之一。

1.4 时间序列分析的应用领域

随着时间序列分析方法的日趋成熟,其应用领域越来越广泛。目前,它不但涉及天文、地理、生物、物理、化学等自然科学领域,而且涉及语音通信、图像识别、雷达声呐、遥感技术、



环境工程、海洋工程等工程技术领域以及国民经济、市场经济、生产管理等社会经济领域。

首先,时间序列分析的应用主要集中在预报预测领域,例如气象预报、市场预测、地震预报、人口预测、汛情预报、产量预测等。其次,是应用于精密测控,例如精密仪器测量、精密机械制造、航空航天轨道跟踪和监控,以及遥控遥测、精细化工控制等。再一个应用领域是安全检测和质量控制。在工程施工和维修中经常会出现异常险情,采用仪表监测和时间序列分析方法可以随时发现问题,及早排除故障,以保证生产安全和质量要求。以上仅仅列举了某些应用领域,实际上还有许多应用,不胜枚举。

我们常常会遇到大量的实测数据,虽然看起来是杂乱无章的,但实际上具有一定的统计规律。很多情况下我们的目的就是找出这种统计规律,进而对其进行深入的研究。时间序列分析是描述动态数据统计特性的一种有效的方法,它是先对动态数据建立数学模型,并在此基础上进一步分析其统计特性。因此模型建立得准确与否至关重要,在建立时间序列模型之前,需要对动态数据进行必要的预处理,剔除不符合统计规律的异常样本,并对样本数据的基本统计特性进行检验,以确保建立时序模型的可靠性和置信度,并满足一定的精度要求。

2.1 平稳性检验

对时间序列进行平稳性检验时,需要考虑两个内容:

- (1) 序列的均值($\overline{x_i}$) 和方差(σ_i^2) 是否为常数。
- (2) 序列的自相关系数(r_i) 是否仅与时间间隔有关,而与时间间隔的端点位置无关。

常用的平稳性检验方法有参数检验法、非参数检验法等几种。

2.1.1 平稳性的参数检验法

平稳性的参数检验法的基本思想是将长序列分成若干个短序列,分别考察各个短序列的均值、方差以及自相关系数之间是否满足平稳性的要求。

假设样本序列 $\{x_t\}$, $t=1, 2, \dots, N$, N 足够大,把该样本序列分成 k 个子序列, $N=kM$, M 也是一个较大的正整数。

$$\left\{ \begin{array}{c} x_{11}, x_{12}, \dots, x_{1M} \\ x_{21}, x_{22}, \dots, x_{2M} \\ \vdots \\ x_{k1}, x_{k2}, \dots, x_{kM} \end{array} \right\} = \{x_{ij}\}$$



$$x_{ij} = x_{(i-1)M+j} \quad (i=1,2,\dots,k; j=1,2,\dots,M)$$

对 k 个样本子序列分别计算它们的样本均值、样本方差和样本自相关系数:

$$\begin{cases} \bar{x}_i = \frac{1}{M} \sum_{j=1}^M x_{ij} \\ \sigma_i^2 = \frac{1}{M} \sum_{j=1}^M (x_{ij} - \bar{x}_i)^2 \\ r_i(\tau) = \frac{1}{M} \sum_{j=1}^M (x_{ij} - \bar{x}_i)(x_{i,j+\tau} - \bar{x}_i) / \sigma_i^2 \end{cases} \quad (i=1,2,\dots,k; \tau=1,2,\dots,m; m \ll M)$$

由平稳性的假定,以上各统计量对不同的子序列 i 不应有显著差异,否则就是否定 $\{x_t\}$ 是平稳序列的假定。设 $\{x_t\}$ 具有理论上的均值 μ 、方差 σ^2 和自相关函数 ρ_τ ,则样本统计量 $\bar{x}_i, \sigma_i^2$ 及 $r_i(\tau)$ 的方差可由随机变量四阶矩的算式得到,如果 k 个子序列的上述统计量没有显著变化,则 $\{x_t\}$ 平稳。

(1) 样本均值的方差

$$\begin{aligned} \sigma_1^2 = D(\bar{x}_i) &= \frac{1}{M^2} E \left[\sum_{j=1}^M \sum_{l=1}^M (x_{ij} - \mu)(x_{il} - \mu) \right] = \\ &= \frac{\sigma^2}{M^2} \sum_{j=1}^M \sum_{l=1}^M \rho_{j-l} = \frac{\sigma^2}{M^2} \left[1 + 2 \sum_{j=1}^M \left(1 - \frac{j}{M}\right) \rho_j \right] \end{aligned}$$

(2) 样本方差的方差

$$\sigma_2^2 = D(\sigma_i^2) = \frac{2\sigma^2}{M^2} \left[1 + 2 \sum_{j=1}^M \left(1 - \frac{j}{M}\right) \rho_j^2 \right]$$

(3) 样本自相关的方差

$$\sigma_3^2(\tau) = D(r_i(\tau)) \approx \frac{1}{M-\tau} \left[1 + \rho_\tau^2 + 2 \sum_{j=1}^{M-\tau} \left(1 - \frac{j}{M-\tau}\right) (\rho_j^2 + \rho_j + 2\rho_{j-\tau}) \right]$$

具体来说,可以采用统计检验法,若取显著水平 $\alpha=0.05$ 和 2σ 原则,此时置信度为 $1-\alpha=0.95$ 。

当

$$|\bar{x}_i - \bar{x}_j| > 1.96 \sqrt{2\sigma_1^2}$$

$$|\sigma_i^2 - \sigma_j^2| > 1.96 \sqrt{2\sigma_2^2}$$

$$|r_i(\tau) - r_j(\tau)| > 1.96 \sqrt{2\sigma_3^2(\tau)} \quad (i \neq j; i, j=1, 2, \dots, k; \tau=1, 2, \dots, m)$$

中的任何一个不等式成立时, $\{x_t\}$ 不具有平稳性。

一般我们并不知道 $\{x_t\}$ 的理论均值、方差与自相关函数,无法直接得到 $\sigma_1^2, \sigma_2^2, \sigma_3^2(\tau)$, 仅能以样本估计值代之。



2.1.2 平稳性的非参数检验法

使用参数检验法判定序列的平稳性时,由于往往得不到序列的理论统计参数,且计算样本统计量的方差时计算也较为复杂,因此实际使用时会受到一定的限制。下面介绍一种实用性很强的平稳性检验方法。

非参数检验法中常使用游程检验法,或称为轮次检验法。在保持随机序列原有顺序的情况下,游程定义为具有相同符号的序列。这种符号可把观测值分成两个相互排斥的类。例如,观测序列的值是 $x_i (i=1,2,\cdots,N)$, 均值为 $\bar{x}$, 若 $x_i \geq \bar{x}$, 用“+”表示, 若 $x_i < \bar{x}$, 则用“-”表示, 按符号“+”“-”的出现顺序将原序列写成如下形式:

$$\underbrace{+++-}_{1} \underbrace{++-}_{2} \underbrace{--}_{3} \underbrace{+-}_{4} \underbrace{--}_{5} \underbrace{--}_{6} \underbrace{+}_{7}$$

“+”“-”共 14 个, 分 7 个游程, 这里不需关注每个游程的长短, 而游程太多或太少则都被认为存在非平稳的趋势。

游程检验所判断的原假设为: 样本的数据出现的顺序为没有明显的趋势, 就是平稳的。

采用的样本统计量: N_1 是一种符号出现的总数, N_2 是另一种符号出现的总数, γ 表示游程的总数, 其中 γ 作为检验统计量。

对于显著水平 $\alpha=0.05$ 时的双边检验, 表 2.1 给出概率分布左右两侧为 $\alpha/2=0.025$ 时的上限 γ_U 和下限 γ_L 。如果 γ 在界限之内, 则接受原假设, 即序列是平稳的, 否则拒绝原假设。

表 2.1 游程检验用 γ 分布表

$N_1 \backslash N_2$		2	3	4	5	6	7	8	9	10	11	12	13	14	15
		γ_L	γ_U	γ_L	γ_U	γ_L	γ_U	γ_L	γ_U	γ_L	γ_U	γ_L	γ_U	γ_L	γ_U
2	γ_L											2	2	2	2
	γ_U														
3	γ_L					2	2	2	2	2	2	2	2	2	2
	γ_U														
4	γ_L				2	2	2	3	3	3	3	3	3	3	3
	γ_U				9	9									
5	γ_L			2	2	3	3	3	3	3	4	4	4	4	4
	γ_U			9	10	10	11	11							
6	γ_L		2	2	3	3	3	3	4	4	4	4	5	5	5
	γ_U			9	10	11	12	12	13	13	13	13			
7	γ_L		2	2	3	3	3	4	4	5	5	5	5	5	6
	γ_U				11	12	13	13	14	14	14	14	15	15	15



续表 2.1

N_1 N_2		2	3	4	5	6	7	8	9	10	11	12	13	14	15
8	γ_L		2	3	3	3	4	4	5	5	5	6	6	6	6
	γ_U				11	12	13	14	14	15	15	16	16	16	16
9	γ_L		2	3	3	4	4	5	5	5	6	6	6	7	7
	γ_U					13	14	14	15	16	16	16	17	17	17
10	γ_L		2	3	3	4	5	5	5	6	6	7	7	7	7
	γ_U					13	14	15	16	16	17	17	18	18	18
11	γ_L		2	3	4	4	5	5	6	6	7	7	7	8	8
	γ_U					13	14	15	15	17	17	18	19	19	19
12	γ_L	2	2	3	4	4	5	6	6	7	7	7	8	8	8
	γ_U					13	14	16	16	17	18	19	19	20	20
13	γ_L	2	2	3	4	5	5	6	6	7	7	8	8	9	9
	γ_U						15	16	17	18	19	19	20	20	21
14	γ_L	2	2	3	4	5	5	6	7	7	8	8	9	9	9
	γ_U						15	16	17	18	19	20	20	21	22
15	γ_L	2	3	3	4	5	6	6	7	7	8	8	9	9	10
	γ_U						15	16	18	18	19	20	21	22	22

例如有 $N=22$ 的观测序列: $++---+-++-++-+-++-$, 因 $N_1=12(+)$, $N_2=10(-)$, $\gamma=11$ 。查表 2.1 得原假设的接受域为 $7 \leq \gamma \leq 17$, 故原序列没有明显的潜在趋势, 是平稳的。

$$P(\gamma \leq \gamma_L) + P(\gamma \geq \gamma_U) = 0.05$$

当 N_1 或 N_2 超过 15 时, 可用正态分布来近似, 即用正态分布表(表 2.2) 来确定检验的接受域和否定域, 此时用的统计量为

$$Z = \frac{\text{游程数} - \text{游程的期望数}}{\text{游程的标准差}} = \frac{\gamma - \mu_\gamma}{\sigma_\gamma}$$

式中

$$\begin{cases} \mu_\gamma = \frac{2N_1N_2}{N} + 1 \\ \sigma_\gamma = \left[\frac{2N_1N_2(2N_1N_2 - N)}{N^2(N-1)} \right]^{1/2} \end{cases} \quad (N = N_1 + N_2)$$

对于 $\alpha=0.05$ 的显著水平, 如果 $|Z| \leq 1.96(2\sigma \text{ 原则})$, 则接受原假设, 否则拒绝。



2.1.3 平稳性的时序图检验法

依据平稳时间序列的均值、方差为常数的特点,可知若某一时间序列为平稳的,其时序图应该表现为该序列值始终在一个常数值附近随机波动且波动范围是有界的这一特点。若该序列的时序图显示出具有明显的趋势项或周期性,则该序列通常不是平稳的。据此可以用来判断一些序列的平稳性,但此方法倾向于定性地描述,使用具有一定的局限性。序列的周期性和趋势项在本章 2.4 节和 2.5 节详细介绍。

2.2 正态性检验

时间序列模型常常建立在具有正态分布特性的白噪声基础上,或者说是从大样本的观点上满足此条件,因此,建立时间序列模型时需要对序列的正态性进行判断。

正态分布的概率密度函数(PDF)记为:

$$p(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp[-(x - \mu)^2 / (2\sigma^2)]$$

式中 μ 和 σ^2 ——样本总体的均值与方差。

概率分布是概率密度函数积分,即

$$P(x < X) = (2\pi\sigma^2)^{-\frac{1}{2}} \int_{-\infty}^X \exp[-(x - \mu)^2 / (2\sigma^2)] dx =$$

$$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{(X-\mu)/\sigma} \exp\left(-\frac{1}{2}x^2\right) dx = \Phi((X - \mu)/\sigma)$$

Φ 称为“概率积分”。标准正态分布函数表见表 2.2。

表 2.2 标准正态分布函数表($\mu = 0, \sigma^2 = 1$)

X	0	1	2	3	4	5	6	7	8	9
-3.0	0.001 3	0.001 0	0.000 7	0.000 5	0.000 3	0.000 2	0.000 2	0.000 1	0.000 1	0.000 0
-2.9	0.001 9	0.001 8	0.001 7	0.001 7	0.001 6	0.001 6	0.001 5	0.001 5	0.001 4	0.001 4
-2.8	0.002 6	0.002 5	0.002 4	0.002 3	0.002 3	0.002 2	0.002 1	0.002 1	0.002 0	0.001 9
-2.7	0.003 5	0.003 4	0.003 3	0.003 2	0.003 1	0.003 0	0.002 9	0.002 8	0.002 7	0.002 6
-2.6	0.004 7	0.004 5	0.004 4	0.004 3	0.004 1	0.004 0	0.003 9	0.003 8	0.003 7	0.003 6
-2.5	0.006 2	0.006 0	0.005 9	0.005 7	0.005 5	0.005 4	0.005 2	0.005 1	0.004 9	0.004 8
-2.4	0.008 2	0.008 0	0.007 8	0.007 5	0.007 3	0.007 1	0.006 9	0.006 8	0.006 6	0.006 4
-2.3	0.010 7	0.010 4	0.010 2	0.009 9	0.009 6	0.009 4	0.009 1	0.008 9	0.008 7	0.008 4
-2.2	0.013 9	0.013 6	0.013 2	0.012 9	0.012 6	0.012 2	0.011 9	0.011 6	0.011 3	0.011 0