·数理统计丛书·

安鸿志 编著

时间序列分析

华东师范大学出版社

61

·数理统计丛书·

时间序列分析

安鸿志　编著

华东师范大学出版社

（沪）新登字第201号

·数理统计丛书·
时间序列分析
安鸿志 编著

华东师范大学出版社出版发行
（上海中山北路3663号）
新华书店上海发行所经销　江苏省句容县排印厂印刷
开本：850×1168 1/32 印张：8.25 字数：210千字
1992年3月第一版　1992年3月第一次印刷
印数：001—2,500本

ISBN7-5617-0702-9/N·020 定价：4.15元

总 序

数理统计是一门应用性很强的学科。它是研究如何有效地收集、整理和分析受随机影响的数据，并对所考虑的问题作出推断或预测，直至为采取决策和行动提供依据和建议的一门学科。凡是有大量数据出现的地方，都要用到数理统计。人口调查、税收预算、测量误差、出生与死亡统计、保险业中赔款额和保险金的确定等，这些数理统计早期主要研究的问题，直到现在仍值得认真研究。建立在现代数学和概率论基础上的数理统计，在近半个世纪以来，在理论、方法、应用上都有较大的发展。抽样调查、试验设计、回归分析与回归诊断、多元分析、时间序列分析、非参数统计、统计决策函数、统计计算、随机模拟、探索性数据分析等统计方法相继产生并在实践中普遍使用，把以描述为主的统计发展到以推断为主的统计。今天，数理统计的内容已异常丰富，应用面广量大，成为当前最活跃的学科之一。

我国科技和经济的发展，需要大量的经过系统训练的数理统计专业人才。最近数年中，国家教育委员会已在一些高校先后设立各种统计专业，这将为我国数理统计发展开创新的局面。为了促进数理统计人才的培养，国家教育委员会于 1984 年召开了"数理统计教学座谈会"，会上交流了各校培养数理统计人才的经验，同时还指出，组织国内专家编写和出版一套数理统计专业的教材是当务之急。

我们认为，一本好的数理统计教材，首先应讲清统计思想与统计方法，所讲的理论应是在实践中有用的统计方法所必需的理论，在阐明各种统计方法时，应给出足够的问题背景和有关的数据，能使学生对数理统计有一个系统、全面的认识，并培养学生对统计实

践的兴趣。另外，文字流畅，带有趣味性，适合大学生阅读，当然也是一本好教材的必要条件。

鉴于国内对数理统计教材的急需，我们约请国内一些颇有造诣的专家，按照上述对教材的特定要求，编写了这套"数理统计丛书"，作为高等院校数理统计专业的基本教材，并将由华东师范大学出版社陆续出版。现在奉献给读者的《时间序列分析》即为该丛书之一。我们希望，这套丛书能对我们数理统计事业的发展起到一定的促进作用，能成为我国年轻一代统计学家的引路之石。

茆 诗 松

1986年11月于华东师范大学

试读结束：需要全本请在线购买： www.ertongbook.com

符 号 说 明

δ_{kj}	克朗内克(Kronecker)记号，即$\delta_{kk}=1$，$\delta_{kj}=0(k\neq j$时)。
A^{τ}	矩阵 A 的转置矩阵。
A^{-1}	满秩方阵 A 的逆矩阵。
$\|A\|$	方阵 A 的行列式,亦用 $\det A$ 表示。
Ex	随机量 x 的均值。
$E(y\|\boldsymbol{x})$	在给定 $\boldsymbol{x}$ 时,y 的条件均值。
$\mathrm{L}(y\|\boldsymbol{x})$	根据随机矢量 $\boldsymbol{x}$ 对 y 做线性最小方差估计。
$\mathrm{Var}(\boldsymbol{x})$	随机矢量 $\boldsymbol{x}$ 的方差阵。
$\mathrm{Cov}(\boldsymbol{x},\boldsymbol{y})$	随机矢量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的协方差阵。
$x\sim N(\mu,\sigma^2)$	随机变量 x 服从正态 $N(\mu,\sigma^2)$分布。
$\boldsymbol{x}\sim N(\boldsymbol{\mu},R)$	随机矢量 $\boldsymbol{x}$ 服从多元正态 $N(\boldsymbol{\mu},R)$ 分布。
χ_s^2	自由度为 s 的 χ^2 分布。
F_{mn}	自由度为 m,n 的 F 分布。
B	后移变换,又称后移算子,如 $Bx_t=x_{t-1}$,$B^k\gamma_j=\gamma_{j-k}$。
∇	差分运算,如 $\nabla x_t=x_t-x_{t-1}$ 。
∇_T	T 步差分运算,如 $\nabla_T x_t=x_t-x_{t-T}$ 。
AR	自回归(Autoregression)。
MA	滑动平均(Moving Average)。
ARMA	自回归滑动平均(Autoregressive Moving Average)。
ARIMA	自回归滑动平均求和(Autoregressive Integrated Moving Average)。
AIC	赤池信息准则 (Akaike's Information Criterion)。
FFT	快速傅里叶变换(Fast Fourier Transform)。

前　言

时间序列分析是概率统计学中的一个比较新的分支。由于它的应用广泛，近年来发展非常迅速，因而它所包括的内容也越来越丰富。从数学角度看，时间序列分析既涉及了随机过程知识，又涉及到数理统计知识。因此，许多熟悉这门知识的人，都是在大学毕业以后才不断了解和掌握它的。近年来，随着概率统计方法的应用越来越普遍，时间序列分析方法也就越来越显得重要，特别是对理工科大学毕业生更是如此。因而有一本比较合适的时间序列分析方面的教科书，也越来越成为迫切的需要了。

时间序列分析的内容丰富，涉及到的基础知识比较多，而希望掌握此门知识的读者又非常广泛。所以，要撰写一本时间序列分析的教科书，并且仅供一学期讲授之用，这不是一件容易的事。况且，在选材方面也有多种不同的方式。作者在完成本书时，曾有以下的宗旨：一是尽量少涉及较深的数学内容，二是尽量多介绍有用的时间序列分析方法，包括较新的概念与方法。作者希望这本书能够成为时间序列分析的入门参考书，或者说是教科书，并希望具有高等数学知识和概率统计基础知识的读者，都能阅读此书。有着不同程度的预备知识的读者，通过阅读本书，对于时间序列分析的基础知识也能有不同层次的了解。但是，作者深知自己的能力有限，此书未必能达到预期的愿望。谨望读者多多赐教和指正。

作　者

1989 年于中国科学院应用数学所

目　录

第一章 引 言

时间序列是指被观测到的依时间次序排列的数据序列。从经济到工程技术，从天文到地理和气象，几乎在各种领域中都会遇到时间序列。对时间序列进行统计分析，简称时间序列分析，这是统计学中的一个重要的分支。

§1 实际背景

我们用以下的实例，说明时间序列的实际背景。

例 1 从 1950 年到 1981 年，我国逐年的水产品总产量，这便是由 32 个数据组成的有次序的数据序列。我们将这一时间序列描绘在图 1.1 中。

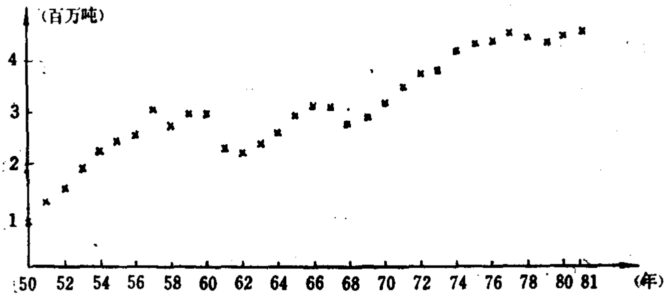

图 1.1 我国历年水产品产量

在经济领域中，类似例 1 的数据序列是大量存在的。比如，历年的工农业总产值，各种商品的逐月产量，等等。

例 2 某地区的逐月降雨量，其实际记录结果，按月份先后排列，便是一个时间序列。为了描绘方便，我们将某地区的连续十年

的逐月降雨量数据序列，在图 1.2 上用曲线连接起来。

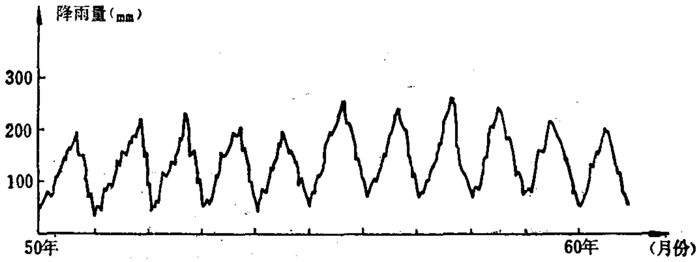

图 1.2　某地区十年中逐月降雨量

气象领域是时间序列遍及的领域。对于很多气象要素的观测结果，都是以时间序列形式出现的。比如，某地的月平均气温、湿度等。

例 3　众所周知，地球自转一周为一天，历时 24 小时。所以，在此例中，时间序列应是以无穷多个 24 构成的数据序列。这样的序列可以称为常值序列。但是，随着现代科学技术对更精确时间量的需要，同时，随着现代测量时间的技术的提高，人们早已发现，地球自转一周所用的时间并不是固定的常数。确切地说，它与 24 小时有极微小的起伏变化量。因此，实际观测到的地球自转一周所用的时间，是一串变化着的数据序列。

例 4　记录某地区各小时的用电平均负荷量，得到一串数据序列，也是一个时间序列。

例 5　某观测站记录到的地下水位的变化过程，是一个随连续时间变化的曲线。为了使用计算机分析这种变化过程，需要对此连续变化过程进行等间隔的采样，采样得到的结果便是一个时间序列。有的观测站，直接按等时间间隔记录地下水位数据，比如每小时记录一次，那么，得到的记录结果也是时间序列。

例 6　用现代医疗器械，可以记录人的两次脉搏之间的时间间隔，这也是时间序列。据医学界认为，此序列包含着关于人体健康状况的丰富信息。

试读结束：需要全本请在线购买： www.ertongbook.com

在医学界，随着各种电子医疗器械的出现，使用时间序列概念越来越普遍。比如，对心电图和脑电图的分析，已经有人在使用时间序列分析方法。

例 7 记录在工作状态下机器产生的振动，再将连续振动过程进行等间隔的离散采样，得到的振动数据也是时间序列。

通过以上几个例子，我们可以看到，时间序列概念的确有着非常广泛的实际背景。

§2 应用分类

在工农业生产中，在自然科学和社会科学的各个领域中，我们会遇到许许多多的时间序列。对于这样诸多的时间序列进行分析，其目的也是多种多样的。在这一节中，我们将时间序列分析的应用背景归纳为以下几类。

1．预报分析　根据对某个变化量的一段观测数据，预报该量在未来时刻的取值，这便是预报问题。比如，在 §1 中的例 1，人们希望通过总结 1950 年到 1981 年的实际水产量的发展规律，预报 1981 年以后的水产量。在分析 §1 中例 2 和例 3 所叙述的时间序列时，也往往是以预报为目的的。将时间序列分析方法用于预报，在时间序列分析的应用中是最广泛的一类应用。本书将在第四章中介绍时间序列的预报方法。

2．控制分析　根据对一个量(或若干个量)的一段观测结果的分析，寻求对某些量的控制措施，以达到某种最优化的目的，这属于控制分析，或称最佳控制设计。在 §1 中的例 4 里，我们记录了过去若干小时的用电负荷数据，其目的在于，通过对这些数据的分析，为供电系统和发电系统提供某种最优化的控制方法。解决这类控制问题，时间序列分析方法是一种重要的数学工具。在近代的最佳控制设计中，使用时间序列分析中的有限参数模型的统计分析方法，是比较常用的方法。关于时间序列分析中的参数模型概念

及其统计方法，将在本书中第二、三两章中分别介绍。至于如何利用时间序列中的参数模型，设计最佳控制方案，则不在本书讨论的范围之内。

3．诊断分析　根据两个不同时间序列的记录值，分析判断它们是否具有相同属性，或者根据一个时间序列的记录值，分析判断它是否具有某些指定的属性，这类分析称为诊断分析，也称为识别诊断。比如，在§1中的例5，我们分析地下水位的数据序列时，如果以预报地震为目的，则希望判断是否出现异常现象。也就是说，根据地下水位的某一段数据序列分析，判断该序列是否仍处于正常状态。类似地，对于§1中的例6而言，医务人员希望根据对一个人的心跳间隔数据序列的分析，判断此人是否健康，这也是要判断一个序列是否处于正常状态。解决这样的判断问题，必然要用到对时间序列的特性的描述方法。这要用到时间序列分析的基础知识，我们将在第二章和第五章中分别介绍时间序列分析中时域和频域方面的基础知识。虽然本书没有专门章节介绍解决上述诊断判别问题的方法，但是，在读者掌握了本书第三章内容之后，自然能够了解解决诊断判别的简单方法。

4．频谱分析　根据时间序列的记录值，分析出此序列中的周期谐波分量，或者对此序列的频率特性进行统计分析，都称为频谱分析。比如，在§1的例7中，由于振动的成因与轴承或传动装置有关，因此，在振动的记录中可能包含有周期分量。找出每个周期分量的频率及其相应的振幅，则要用到时间序列分析中的谱分析方法。除了寻找谐波分量外，还要对非谐波分量的频率分布情况进行统计分析。这些内容同样属于时间序列分析中的谱分析。我们将在第五章中介绍时间序列的谱域描述方法，在第六章中介绍时间序列的谱分析方法。

在实际应用中，经常要涉及多元时间序列。所谓多元时间序列，是指同时观测到的多个数据序列，或者称为多元数据序列。对于多个数据序列进行联合的统计分析，称为多元时间序列分析。在

多元时间序列分析的应用中，也有上述的四种不同的类型。本书以介绍一元时间序列分析为主，仅在第七章中，概述多元时间序列分析中的某些主要内容。

第二章　时间序列的时域描述方法

为了从实际的时间数据序列中提取有用的信息，或者说，从随机数据序列中分析出相对稳定的统计性规律，我们需要掌握时间序列的概率定义和各种性质。在这一章里，我们将介绍时间序列的概率定义以及在时间域里的各种重要的描述方法。

§1　随机过程与时间序列

一、随机过程

现实世界中的大量的物理过程，都或多或少地卷入一些随机性的因素，对这样的物理过程的量测结果，用随机过程描述是最合适的。它描述了在时间结构上遵从一定概率分布的数量变化过程。

在数学上，随机过程被定义为一簇随机变量，即 $\{x(t), t\in T\}$，其中 T 表示时间 t 的变动范围，对每个固定的时刻 t 而言，$x(t)$ 是一个普通的随机变量，这些随机变量的全体，就构成一个随机过程。比如，当 $T=(-\infty,\infty)$ 时，随机过程可以表示成 $\{x(t), -\infty<t<\infty\}$，其中 $x(t)$ 是时间 t 的随机函数，因为在每个时刻 t，$x(t)$ 为一个随机变量。

当 $T=\{0,\pm1,\pm2,\cdots\}$ 时，也就是说，时刻 t 只取整数值时，随机过程 $\{x(t), t\in T\}$ 则可写成如下形式，$\{x(t),\ t=0,\ \pm1,\ \pm2,\cdots\}$。此类随机过程 $x(t)$，是离散时间 t 的随机函数，所以也称它为随机序列。由于 t 代表时间，所以此类随机序列也称为时间序列。比如，对于一个连续时间的随机过程的等间隔采样序列，即 $\{x(th),\ t=0,\ \pm1,\ \pm2,\cdots\}$，就是一个随机序列。

我们将要讨论的时间序列分析，只是对上述随机序列的一部分特殊类型的统计分析，而不是对其所有的随机序列进行统计分析。粗略地说，时间序列分析所讨论的随机序列，主要包括平稳序列及其有关的随机序列。所以，本书主要涉及的随机序列，多与平稳序列有关。我们将在下一节中介绍平稳序列概念。在这一节中，我们先介绍一般时间序列的分布、均值与协方差等概念。

二、时间序列的分布、均值与自协方差函数

1．时间序列的概率分布

一个时间序列便是一个无限维的随机矢量。它的概率分布是有限维随机矢量概率分布的推广。一个有限维随机矢量 $\boldsymbol{x}=(x_1,x_2\cdots,x_p)^{\tau}$，它的分布是一个 p 维的概率分布。那么，一个无限维的随机矢量 $\boldsymbol{x}=(\cdots,x(-1),x(0),x(1),\cdots)^{\tau}$，其概率分布应当用一个无限维概率分布描述。但是，根据柯尔莫哥洛夫定理(见〔6〕第三册第一章§2)，一个时间序列的概率分布，可以用它的有限维分布簇来描述。例如，时间序列的所有的一维分布是：

$$\cdots,\ F_{-1}(\cdot),\ F_0(\cdot),\ F_1(\cdot),\ \cdots$$

这无穷多个一维分布，称为时间序列的一维分布簇，其中每一个分布 $F_k(\cdot)$，表示 $x(k)$ 的分布函数。又比如，它的所有可能的二维分布是：

$$F_{i,j}(\cdot,\ \cdot),\ i,\ j=0,\ \pm1,\ \pm2,\ \cdots,\ i\neq j$$

这些二维分布函数的全体，称为时间序列的二维分布簇，其中每个分布 $F_{ij}(\cdot,\cdot)(i\neq j)$，表示 $(x(i),\ x(j))$的二元随机变量的联合分布。一个时间序列的所有有限维分布簇的全体，称为该序列的有限维分布簇。

一个随机变量的概率结构，被它的分布函数唯一决定。一个时间序列的概率结构，被它的有限维分布簇唯一决定。在统计学中，一个随机变量的分布函数，决定了随机变量的全部统计特征。但是，用分布函数描述随机变量的特征，有其不方便之处，所以常用随机变量的某些重要数字特征来描述随机变量的性质。比如常用

随机变量的均值、方差来描述随机变量的重要特征。这些特征量虽然不是随机变量的全部特征的描述，但是，它们往往能代表随机变量的主要概率特征，而且它们易于计算和在实际中使用。类似地，在时间序列分析中，由于分布函数簇的复杂性，人们更加注意使用时间序列的各种数值特征量的描述。比如即将引入的均值函数，协方差函数，以及在第五章中引入的谱分布函数等。所谓时间序列的统计分析，主要是对这些特征量的统计分析，很少涉及对有限维分布簇的统计分析。

2．均值函数

一个时间序列 $\{x(t), t=0, \pm 1, \pm 2, \cdots,\}$，我们将用以下的简化记号代替，即 $\{x_t\}$，其中 $x_t=x(t)$，$t=0$，± 1，± 2，…。所谓时间序列 $\{x_t\}$ 的均值函数，系指

$$\mu_t=Ex_t=\int_{-\infty}^{\infty} x dF_t(x), \tag{2.1.1}$$

其中 Ex_t 表示在 t 固定时对随机变量 x_t 求均值，它只与一维分布簇中的分布函数 $F_t(\cdot)$ 有关。当 t 取遍所有可能的整数值时，形成了离散时间的函数 μ_t，称 μ_t 为 $\{x_t\}$ 的均值函数。实质上，μ_t 只是一个实数列，它被 $\{x_t\}$ 的一维分布簇所决定。

3．协方差与自相关函数

时间序列的协方差函数与随机变量之间的协方差相似，它被定义为

$$\begin{aligned}\gamma(t, s)&=E(x_t-\mu_t)(x_s-\mu_s)\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_t)(y-\mu_s)dF_{t,s}(x, y),\end{aligned} \tag{2.1.2}$$

式中 $F_{t,s}(x,y)$ 为 (x_t, x_s) 的二维联合分布。当 t,s 取遍所有可能的整数值时，形成了二元离散时间函数 $\gamma(t,s)$，称 $\gamma(t,s)$ 为时间序列 $\{x_t\}$ 的自协方差函数。它被 $\{x_t\}$ 的二维和一维分布簇所确定。对于固定的 t 和 s，x_t 和 x_s 是两个随机变量，它们的协方差为

$$\mathrm{Cov}(x_t, x_s)=E(x_t-\mu_t)(x_s-\mu_s)=\gamma(t, s).$$

试读结束：需要全本请在线购买： www.ertongbook.com