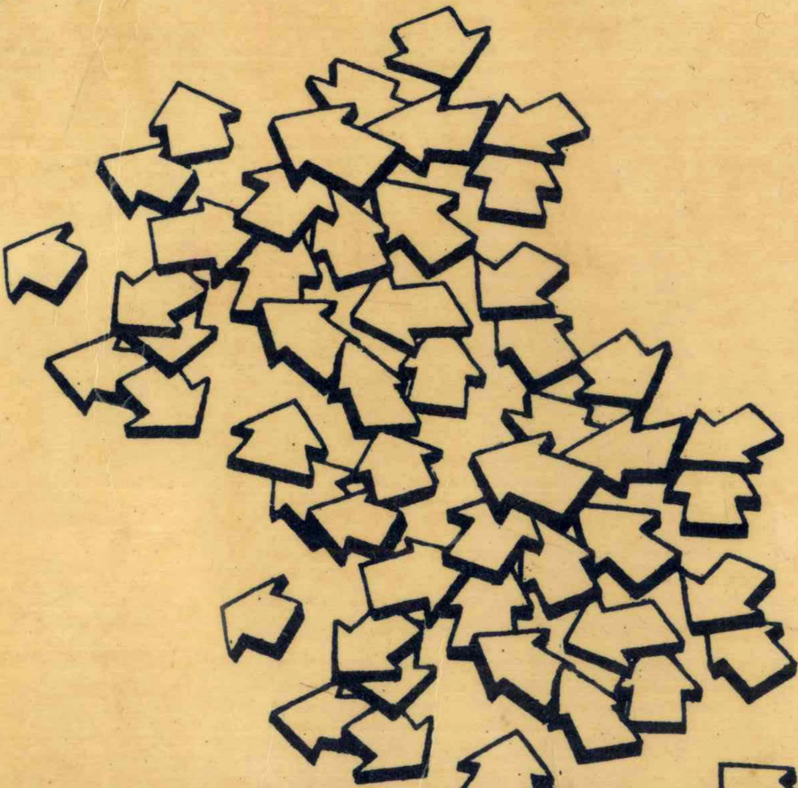
时间序列分析

主　编　杜丹青
副主编　黎　实

西南财经大学出版社

SHIJIAN XULIE FENXI
ZHUBIAN　DU DANQING
FUZHUBIAN　LISHI
XINAN CAIJING DAXUE
CHUBANSHE



时间序列分析

主　编　杜丹青
副主编　黎　实

西南财经大学出版社

（川）新登字017号

责任编辑：傅　虹
封面设计：梁建成

书　名：时间序列分析
主　编：杜丹青
副主编：黎　实

出　版：西南财经大学出版社
（四川省成都市光华村西南财经大学内）
邮编：610074　电话：（028）7763785
印　刷：四川气象印刷厂
发　行：西南财经大学出版社
四川省新华书店经销
开　本：787×1092mm　1/32
印　张：8
字　数：165千字
版　次：1995年11月第1版
印　次：1995年11月第1次印刷
印　数：800册
定　价：12.80元

ISBN7—81017—983—7/F・810

前　言

时间序列分析作为统计数学的一个分支，是分析动态数据、提供定量预测的重要理论和方法。时间序列分析理论可分为“频域分析”和“时域分析”两大类。前者涉及平稳序列及其相关函数的谱理论和富里埃变换的有关知识，而后者则借助于前者的若干理论研究结果，更多地通过对序列在时间域上的各种平均值的分析和研究，揭示动态数据所遵从的统计规律性。自本世纪 70 年代，G. E. P. Box 和 G. M. Jenkins 的著作［1］问世以来，“时域分析”方法日益显示其科学性和实用性，而受到更多的重视。

以 Box—Jenkins 命名的现代时间序列分析方法，将预测理论与计算技术结合起来，为统计预测方法在生产、科技和社会经济各方面的广泛应用创造了有利条件。

经济系统是一个随时间演变的复杂动态系统。经济变量通常表现为按时间顺序排列，彼此之间呈现出相互影响、相互制约而且又与它们的历史状况密切关联、错综复杂的关系。以统计方法测定发生于经济生活中的具体数量规律性，除熟知的计量经济学和计量经济模型外，现在也越来越多地使用时间序列分析方法和时间序列模型。但它们之间有着一些重要的区别：计量经济学必须依据一定的经济理论建立经济模型，其建模的目的是为了检验经济理论，寻求经济规律进而为经济决策提供依据，而时间序列分析方法较多的则不去区分引起经济现象演变发展的种种原因，而只是将经济变量的

现在值看作它的过去值和随机干扰的函数，并在一定的数学模型假定之下，利用变量自身的历史资料，寻求观测数据演变所遵从的动态规律性，以预测其未来。建立时间序列模型的目的主要是为了预测。由于时间序列分析方法是处理动态数据的一种探索性方法，它不需要以任何经济理论为根据，因此通常不将它归结在计量经济学范畴之内。但值得注意的是：对经济变量动态特性和周期特性的研究；经济变量之间因果关系的测定和检验；联立方程建模等许多方面新的理论发展，已经将时间序列分析方法和标准的计量经济学方法密切地联系起来，成为计量经济学的一个重要组成部分了。本书第六、七两章所介绍的关于传递函数模型和干预影响分析的内容能较好地说明这一点。

B—J 时间序列分析方法在理论上比较完善，预测精度高，在专用软件的支持下，使用也十分方便，现已在宏观经济分析、市场预测、人口研究、汇率分析等许多方面取得不少的应用成果。为满足经济类院校的教学需要，以较小的篇幅，避免一些较为艰深的理论分析和数学推导，而又能较为完整地了解到 B—J 方法的基本面貌，并能借助适当统计软件的具体使用，学会对时间序列数据建模的反复探索过程，选择较佳模型并进行实际预测分析，我们根据在西南财经大学为统计系研究生和本科学生开设相应课程所用讲义，并结合多年教学和研究工作的一些体会、心得编写此书，以供经济类院校师生和从事统计分析和定量预测的实际工作者阅读参考。

本书共八章，前六章由杜丹青执笔，第七章由杨峰执笔，第八章由黎实执笔，贾栩参加实例计算。全书由杜丹青、黎

实总纂定稿并任主编、副主编。

本书的出版，列入学校教材出版资助，并得到西南财经大学出版社和统计系的大力支持，在此深表谢意。

限于作者水平，书中不当之处在所难免，恳请读者批评指正。

编　者

1995年1月

目　录

第一章　引　论

§1.1　经济预测与时间序列分析

预测和决策问题在社会经济活动中广泛存在。如宏观经济调控、资源合理配置、生产优化、市场需求和经济结构分析等等。预测是对未来的陈述，决策则是对未来行动的部署。二者紧密联系，交错而又难以分割，但科学、合理的决策必须建立在准确而可靠的预测基础之上。

经济预测是对经济现象未来发展变化的推测，其目的在于确定合理的经济结构和体制，为制定经济政策和经济规划提供可靠的依据，故其意义十分重要。由于经济系统庞大、复杂、多层次、随时间演变而不可逆转以及人们自身的参与而受主观心理因素的影响等原因，经济预测问题难度很大且精度不高，许多方面均有待不断探索和改进。

经济预测可按不同方式而加以分类。为加深对预测方法的了解，下面就从几种预测类型简略加以介绍：

1. 主观预测（Subjective Forecasting）和特尔裴法（Delphi Method）　在本世纪 40 年代首先为美国兰德公司（Rand Corporation）使用。

2. 结构和计量经济模型　用一族数学函数表达经济现象中复杂的因果关系。例如在完全市场条件下，第 t 年商品的价格 P_t 由市场的供需关系确定。设 W_t 为生产该商品的工时工资，Y_t 为人均收入，Q_t^s 和 Q_t^d 分别表示对该商品的供给量和

需求量，则有

$$Q_t^s=f_s(P_t,W_t) \quad Q_t^d=f_d(P_t,Y_t) \tag{1.1}$$

特别，取 f_s 和 f_d 为线性函数并考虑其他影响因素而引入扰动项，可将（1. 1）式修改为：

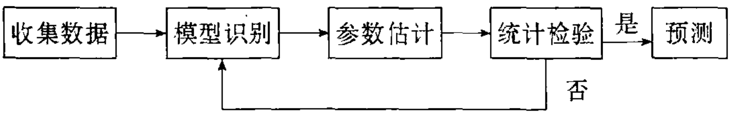
$$Q_t^s=\alpha_0+\alpha_1 P_t+\alpha_2 W_t+\varepsilon_t'$$

$$Q_t^d=\beta_0+\beta_1 P_t+\beta_2 Y_t+\varepsilon_t'' \tag{1.2}$$

其中 ε_t'，ε_t'' 可考虑为具有零均值的某类随机变量。(1. 2) 成为一种统计模型 (Statistical Model) 或计量经济模型 (Econometric Model)。模型中未知参数 α_i，β_i 有待估计，可按某种原则加以确定。由经济学中均衡或调节方程（如假定市场商品当年售完 $Q_t^s=Q_t^d$）可得出关系式

$$P_t=a_0+a_1 W_t+a_2 Y_t+e_t \tag{1.3}$$

即可通过对 W_t，Y_t 和 e_t 的预测和估计作出对未来均衡价格 P_t 的预测。这种预测的误差可能来自不同的方面：对 W_t，Y_t 的预测误差；扰动项 e_t；参数估计带来的误差和模型误差。当 P_t 的预测误差超过一定的限度时，预测将失去现实意义，因此要考虑预测模型的科学性和可检验性（基于科学的理论和方法，可靠的数据，先进的计算技术并制定相应的检验法则）。其研究过程可用以下的流程图表示：



3. 时间序列分析（动态数据处理） 时间序列数据指一批按时间先后顺序记录下来的观测结果：Y_1，Y_2，…，Y_t…

（例如某地区一种商品的月销售量），其发展变化受到一系列其他因素的影响（如商品价格、商品质量、居民人数、收入水平、商业管理水平、消费心理、政策调整甚至气候条件、自然灾害等等），而 Y_t 则可视为这些因素共同影响的结果。时间序列分析方法不去区分（事实上，也很难区分）这些影响因素是如何在起作用，而只是从它们综合影响的结果，即事物自身发展演变的情况，去寻求其客观变化的规律性，并在这些影响因素大致保持稳定的前提下，去推断 Y_t 的未来发展。

由于时序分析方法只需要利用到 $\{Y_t\}$ 自身的历史数据，它是十分经济的；当引起事物发展变化的外部条件未发生显著变化时（特别对于短期预测而言），它是既经济又有效的。

通常作为时序变量 Y_t 有如下的分解形式：

（1）加法模型：$Y_t = T_t + S_t + C_t + I_t$　　（1. 4）

（2）乘法模型：$Y_t = T_t \cdot S_t \cdot C_t \cdot I_t$　　（1. 5）

其中 T_t（Trend）为长期趋势部分。表现为按某种确定规律，稳步增减变化。通常可用 t 的线性函数、指数函数或多项式描述。

S_t（Seasonal）为季节性周期部分。表现为具有固定周期的规律变化，又称商业循环（Business Cycle）。通常可用三角函数、指数振幅三角函数或季节指数描述。

C_t（Cycle）为循环变化，指周期不固定的波动起伏。

I_t（Irreqular）为不规则变化，由许多不可控因素引起，也称残量变化。

在加法模型中，各分量有相同的计量单位，而在乘法模型中，除 Y_t 与 T_t 单位相同外，其余各分量则是相对于趋势分量的比例，故又称比例模型。在 Y_t 的分解形式中，前两部分

称为时间序列 $\{Y_t\}$ 的确定性部分。它们在每一时刻 t 的值可以用 t 的某个确定性函数表示。后两部分统称时间序列的随机性部分，对于每一固定的 t，它是一个随机变量。

4. 确定性模型　在实际应用中，如果时间序列的随机性部分相对说来并不显著，或者对分析结论的精确性要求不很高，这时可将序列看作确定性的。Y_t 基本上可用 t 的确定性函数表示：

$$Y_t=f(t)+\varepsilon_t \tag{1.6}$$

其中 $\{\varepsilon_t\}$ 为误差序列，常假定为零均值、等方差且互不相关的随机变量序列。

根据具体情况，f (t) 可以适当选择为 t 的多项式、三角函数或具有指数振幅的三角函数。而模型中的未知参数按照某种优良性准则加以确定。

在社会经济领域，变量所形成的时间序列大多呈现某种增长趋势。基于一定的理论分析，也可选择适当的理论曲线作为拟合曲线。例如：

指数曲线　$Y_t=ae^{bt}$

修正指数曲线　$Y_t=K+ae^{bt}$

龚柏兹（Compertz）曲线　$Y_t=Ka^{b^t}$

逻辑斯蒂（Logistic）曲线　$Y_t=K\left[1+ae^{f(t)}\right]^{-1}$

其中 K、a、b 为满足一定条件的未知参数，f (t) 是 t 的多项式，可根据相应曲线的特性，对其中未知参数加以估计。有关具体内容可参阅经济预测方面的著作。

5. 某些特定的预测公式

设 $\{Y_t\}$ 为时间序列，用 $\hat{Y}_m(l)=f_l(Y_1, \cdots, Y_m)$ $(l \geq 1)$ 表示基于 $\{Y_t, t \leq m\}$ 对 Y_{m+l} 的预测，其中 f_l 是与 l 有关

的 $\{Y_t, t\leqslant m\}$ 的函数，l 为预测步长。下面我们根据一些"朴素的"直观想法，给出几个预测公式。

$$1°\quad \hat{Y}_m(1)=Y_m \tag{1.7}$$

考虑一个简单的例子，$\{Y_t\}$ 满足链状关系

$$Y_t=Y_{t-1}+\varepsilon_t \tag{1.8}$$

$\{\varepsilon_t\}$ 是零均值、独立同分布的随机变量序列。特别当 $P\{\varepsilon_t=1\}=P\{\varepsilon_t=-1\}=\frac{1}{2}$，$\{Y_t\}$ 称为对称的简单随机游动。

设想从初始状态 $Y_0=i$ 出发，序列 $\{Y_t\}$ 可表为：

$$\begin{aligned}
&Y_1=Y_0+\varepsilon_1\\
&Y_2=Y_0+\varepsilon_1+\varepsilon_2\\
&\cdots\cdots\cdots\cdots\\
&Y_t=Y_0+\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_t
\end{aligned} \tag{1.9}$$

已知 $Y_1, \cdots, Y_m$ 时，对 Y_{m+l} 所作的预测 $\hat{Y}_m(l)$ 通常应满足某种优良性条件，例如使预测均方误差 $E[(Y_{m+l}-\hat{Y}_m(l))^2]$ 达到最小（这时 $\hat{Y}_m(l)$ 称为 Y_{m+l} 的最小均方误差预测），可证：我们有

$$\hat{Y}_m(l)=E[Y_{m+l}|Y_1,\cdots,Y_m] \tag{1.10}$$

即 Y_{m+l} 关于 $Y_1, \cdots, Y_m$ 的条件均值。

利用条件均值的性质和对误差序列 $\{\varepsilon_t\}$ 所作假定，我们有

$$\begin{aligned}
E[Y_{m+l}|Y_1,\cdots,Y_m]&=E[Y_m+\varepsilon_{m+1}+\cdots+\varepsilon_{m+l}|Y_1,\cdots,Y_m]\\
&=Y_m
\end{aligned} \tag{1.11}$$

并可得出预测误差方差

$$\mathrm{Var}[Y_{m+l}|Y_1,\cdots,Y_m]=l\sigma^2 \tag{1.12}$$

其中 $\sigma^2=\mathrm{Var}(\varepsilon_t)$，在正态假定下，可以得出对 Y_{m+l} 的最小均方误差预测是 $\hat{Y}_m(l)=Y_m$，而 Y_{m+l} 的置信系数为 95% 的置信区间为：

$$(\hat{Y}_m(l)-1.96\sqrt{l}\,\sigma,\quad \hat{Y}_m(l)+1.96\sqrt{l}\,\sigma)$$

2° $\hat{Y}_m(1)=Y_m+(Y_m-Y_{m-1})=2Y_m-Y_{m-1}$，即保持前一期的改变量。

3° 滑动平均(Moving Average)

$$\hat{Y}_m(1)=\frac{1}{n}(Y_m+Y_{m-1}+\cdots+Y_{m-n+1}) \tag{1. 13}$$

其中 n 是滑动平均项数。

滑动平均的特点是“修匀”原序列中的某些不规则变化，例如特大值、特小值，使之平滑化，并使趋势倾向更加明显。当原序列无趋势变化或趋势变化很小时，滑动平均可直接用于预测。当原序列存在线性趋势或其他趋势变化时，还可对序列采用二次或三次滑动平均。总之选用滑动平均之目的，在于消除原序列中不规则变化之影响，显现出序列所遵从的基本模式，然后有针对性地得出适当的预测公式。滑动平均不足之处是对前 n 期数据等同看待，且滑动平均的项数 n 如何选择也是一个有待考虑的问题。

4° 指数修匀模型 (Exponential Smoothing Model) 它是对滑动平均的一种改进。事实上，在滑动平均预测公式中我们已作如下假定：

(1) 下一期的预测值只与前 n 期的数值有关。

(2) 前 n 期的数值对预测值的影响是相同的。

前一项假定使得我们不能充分利用数据所带来的信息，而后一项假定与实际情况不相符合，因为距离预测期愈远的

数据．对预测所起的作用将愈小。为此我们对滑动平均预测公式作如下改进：由（1．13）我们有

$$\hat{Y}_{m-1}(1)=\frac{1}{n}(Y_{m-1}+Y_{m-2}+\cdots+Y_{m-n})$$

和

$$\hat{Y}_m(1)=\hat{Y}_{m-1}(1)-\frac{1}{n}Y_{m-n}+\frac{1}{n}Y_m \tag{1.14}$$

用 $\hat{Y}_{m-1}$（1）代替远期值 Y_{m-n}，并将 $\frac{1}{n}$ 改为 β．$0<\beta<1$，则（1．14）变成

$$\hat{Y}_m(1)=\beta Y_m+(1-\beta)\hat{Y}_{m-1}(1) \tag{1.15}$$

$$=\hat{Y}_{m-1}(1)+\beta(Y_m-\hat{Y}_{m-1}(1)) \tag{1.16}$$

（1．15）表明对 Y_{m+1} 的预测值 $\hat{Y}_m$（1）是前一期（$t=m-1$）对 Y_m 的预测值 $\hat{Y}_{m-1}(1)$ 和 $t=m$ 时的实测值 Y_m 的加权平均（权系数分别为 $1-\beta$ 和 β)。而（1．16）可理解为对 Y_{m+1} 的预测值 $\hat{Y}_m$（1）是前一期对 Y_m 的预测值 $\hat{Y}_{m-1}$（1）与一个修正值β（$Y_m-\hat{Y}_{m-1}$（1））之和。其中 $Y_m-\hat{Y}_{m-1}$（1）为前一期的预测误差，β 为修正系数。故（1．16）表明预测公式可以通过新获取的信息 $Y_m-\hat{Y}_{m-1}(1)$ 而得到修正。这是一种自适应预测，且存储数据最少。

由公式（1．15）可递推得到

$$\begin{aligned}\hat{Y}_m(1)&=\beta Y_m+(1-\beta)\hat{Y}_{m-1}(1)\\&=\beta Y_m+(1-\beta)(\beta Y_{m-1}+(1-\beta)\hat{Y}_{m-2}(1))\\&=\beta Y_m+\beta(1-\beta)Y_{m-1}+(1-\beta)^2\hat{Y}_{m-2}(1)=\cdots\cdots\\&=\beta Y_m+\beta(1-\beta)Y_{m-1}+\beta(1-\beta)^2Y_{m-2}+\cdots\\&\quad+\beta(1-\beta)^kY_{m-k}+\cdots\end{aligned} \tag{1.17}$$

即预测值可表为历史值的加权平均，权数系数之和为：