第五章　非平稳序列的随机分析

非平稳序列的分析

两种不同的角度

一个方向是由外向内的分析视角

确定性因素分解方法

一个方向是由内向外的分析视角

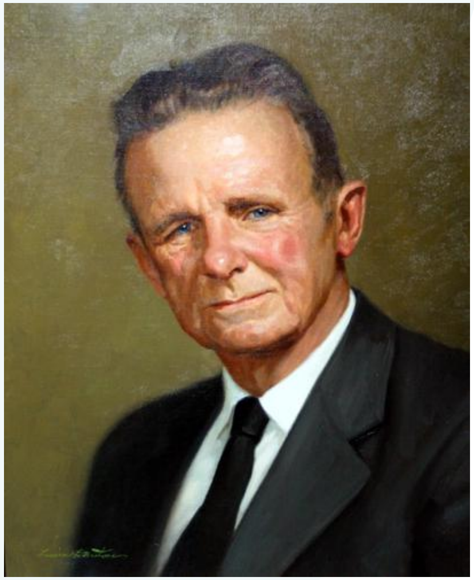
ARIMA方法

本章结构

1. Cramer分解
2. 差分运算
3. 单位根检验
4. ARIMA模型

Cramer分解定理

- Cramer分解定理
  - Harald Cramer, （1893-1985)。瑞典人，斯德哥尔摩大学教授，Wold的指导教师，著名的统计学家和保险精算学家。
  - Cramer 分解定理是Wold分解定理的理论推广，它说明任何一个序列的波动都可以视为同时受到了确定性影响和随机性影响的综合作用。平稳序列要求这两方面的影响都是稳定的，而非平稳序列产生的机理就在于它所受到的这两方面的影响至少有一方面是不稳定的。



Cramer分解定理（1961）

- 任何一个时间序列$\{x_t\}$都可以分解为两部分的叠加：其中一部分是由多项式决定的确定性趋势成分，另一部分是平稳的零均值误差成分，即

$$x_t = \mu_t + \varepsilon_t$$

$$\mu_t = \sum_{j=0}^{d} \beta_j t^j$$

确定性影响

$$\varepsilon_t = \Psi(B)a_t$$

随机性影响

本章结构

1. Cramer分解
2. 差分运算
3. 单位根检验
4. ARIMA模型

差分运算

- 一阶差分

$$\nabla x_t = x_t - x_{t-1}$$

- p阶差分

$$\nabla^p x_t = \nabla^{p-1} x_t - \nabla^{p-1} x_{t-1}$$

- k步差分

$$\nabla_k x_t = x_t - x_{t-k}$$

差分运算的实质

- 差分方法是一种非常简便、有效的确定性信息提取方法
- Cramer分解定理在理论上保证了适当阶数的差分一定可以充分提取确定性信息
- 差分运算的实质是使用自回归的方式提取确定性信息

$$\nabla^d x_t = (1-B)^d x_t = \sum_{i=0}^{d} (-1)^i C_d^i x_{t-i}$$

差分方式的选择

- 序列蕴含着显著的线性趋势，一阶差分就可以实现趋势平稳
- 序列蕴含着曲线趋势，通常低阶（二阶或三阶）差分就可以提取出曲线趋势的影响
- 对于蕴含着固定周期的序列进行步长为周期长度的差分运算，通常可以较好地提取周期信息

例5.1

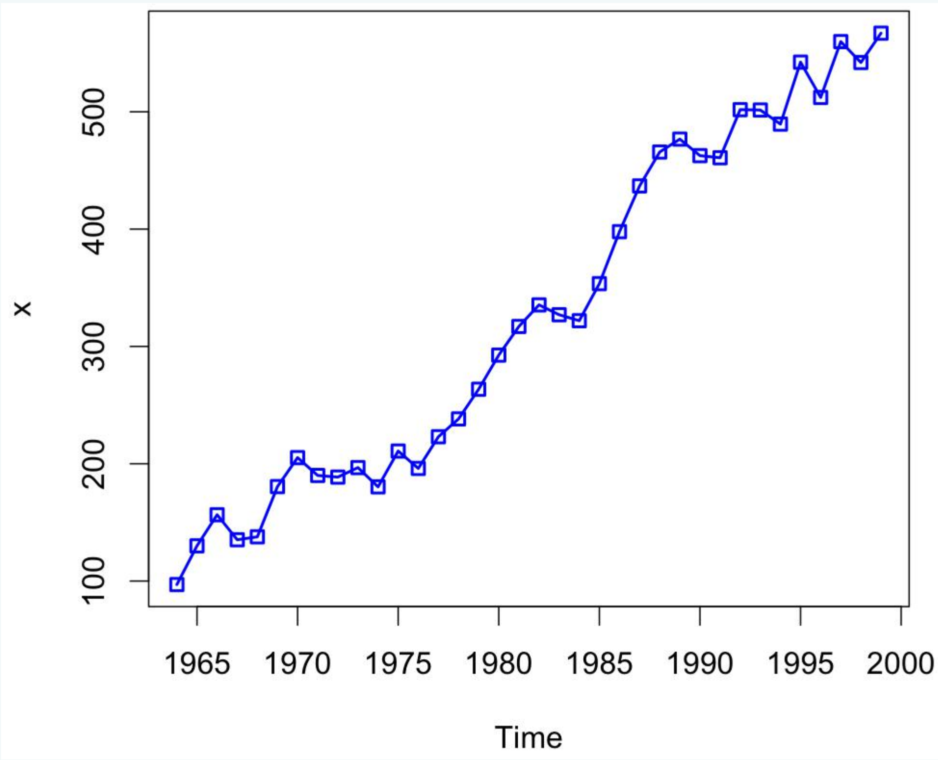
【例5.1】1964年——1999年中国纱年产量序列蕴含着一个近似线性的递增趋势。对该序列进行一阶差分运算

考察差分运算对该序列线性趋势信息的提取作用

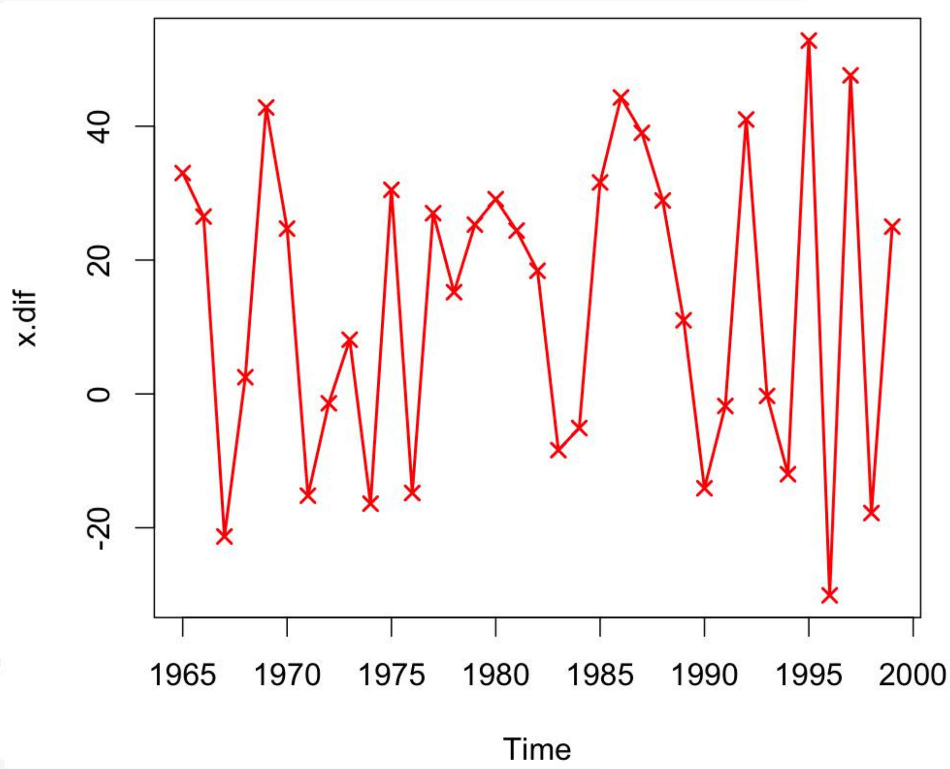
$$\nabla x_t = x_t - x_{t-1}$$

差分前后时序图

❖ 原序列时序图

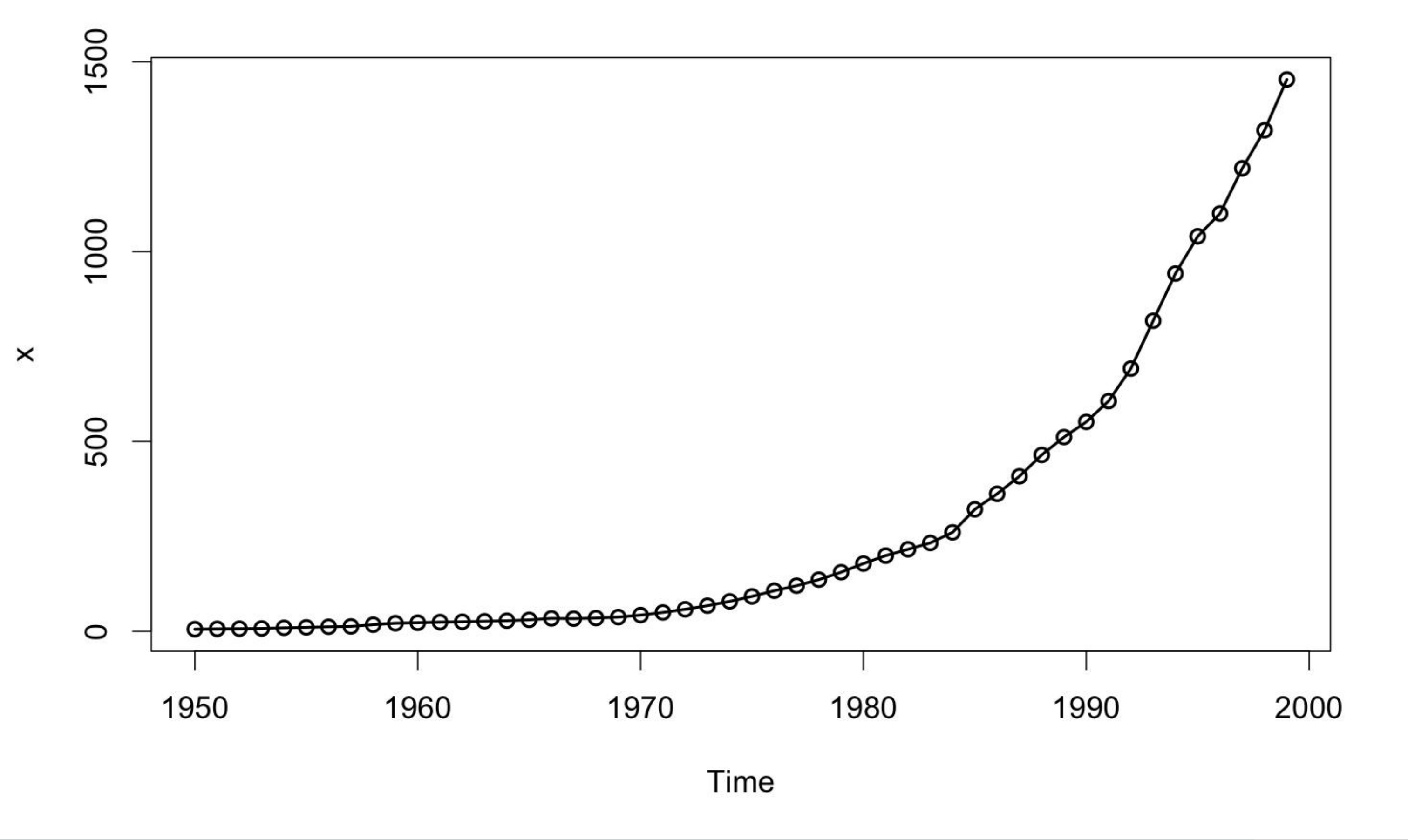


❖ 差分后序列时序图



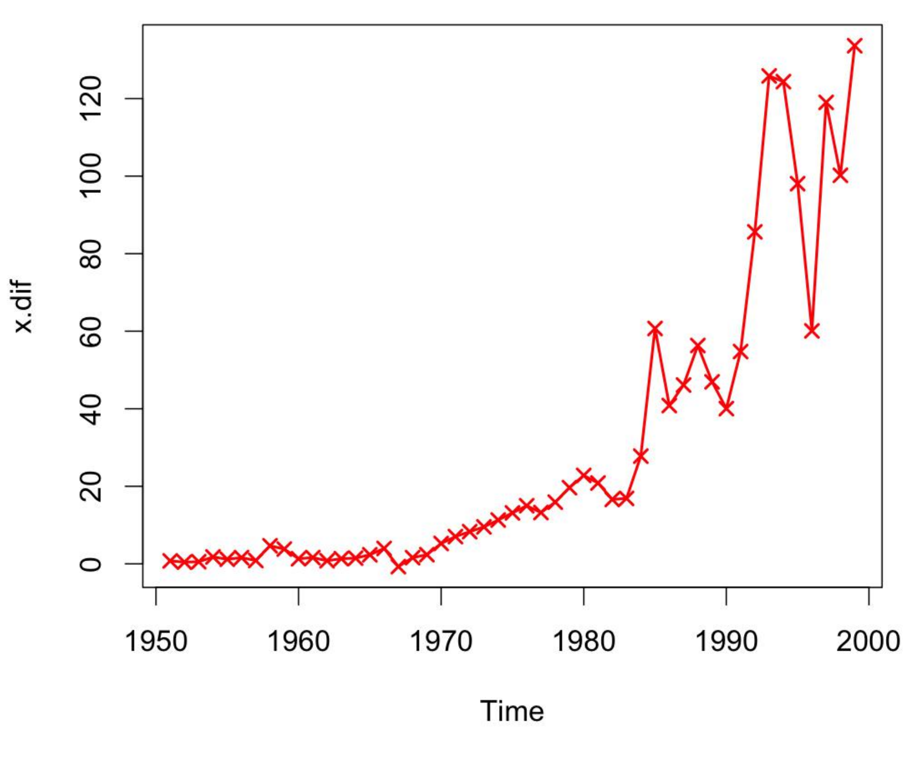
例5.2

- 尝试提取1950年——1999年北京市民用车辆拥有量序列的确定性信息

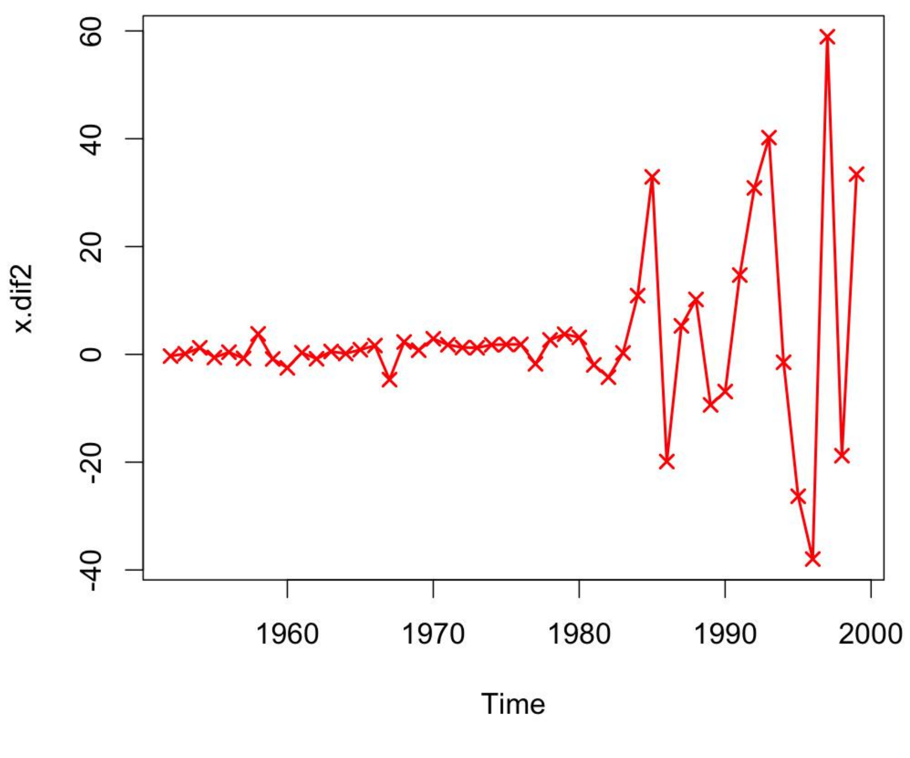


差分后序列时序图

❖ 一阶差分

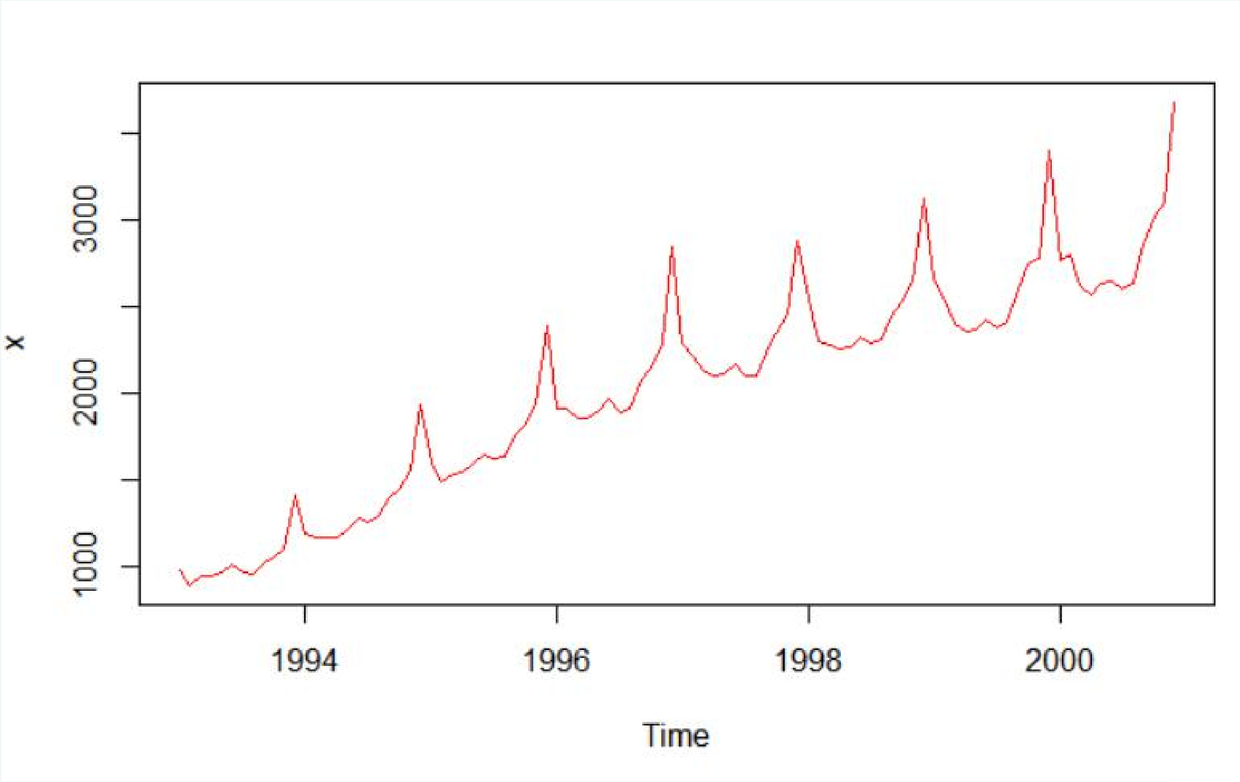


❖ 二阶差分



例5.3

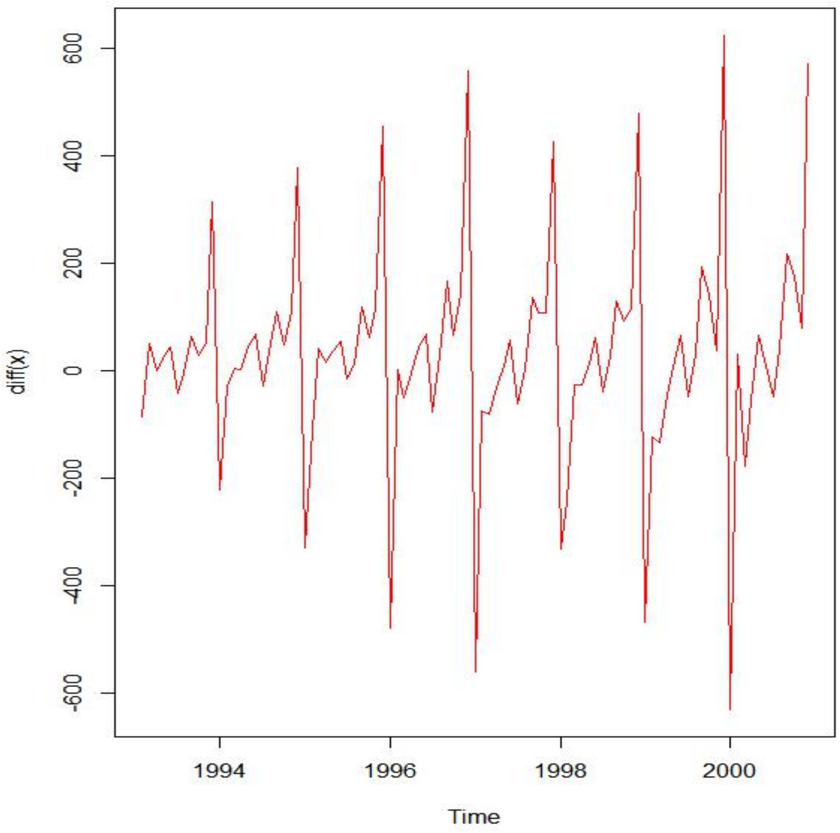
- 差分运算提取1993-2000年中国社会消费品零售总额序列序列中的确定性信息

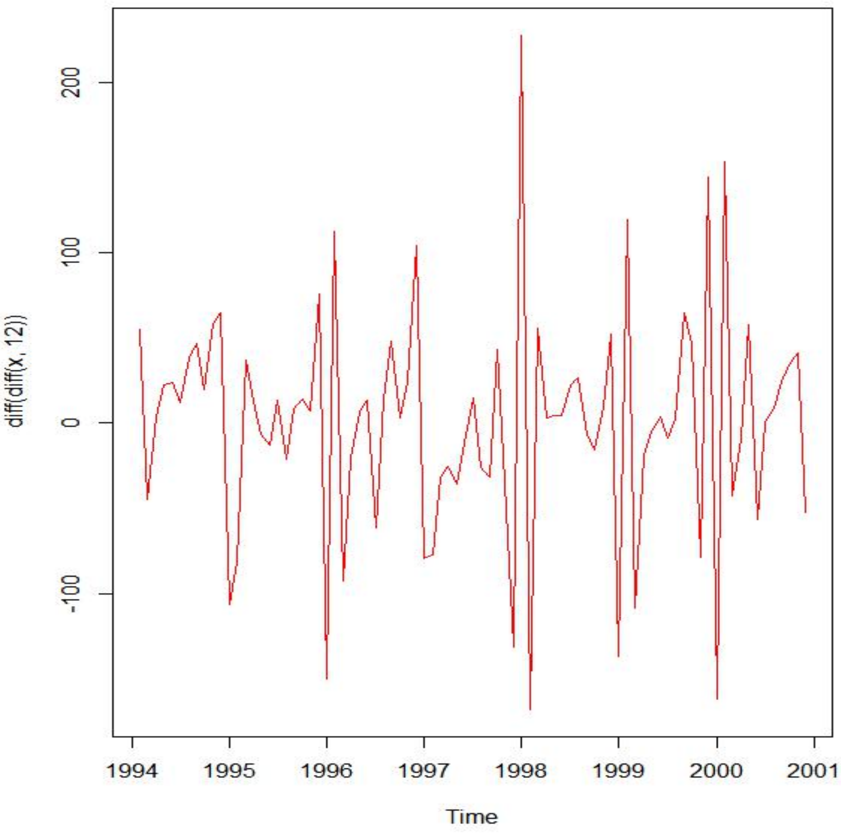


差分后序列时序图

❖ 一阶差分

❖ 1阶－12步差分





过差分

- 从理论上而言，足够多次的差分运算可以充分地提取原序列中的非平稳确定性信息。
- 但应当注意的是，差分运算的阶数并不是越多越好。因为差分运算是一种对信息的提取、加工过程，每次差分都会有信息的损失。
- 在实际应用中差分运算的阶数得适当，应当避免过度差分的现象 。

例5.4

- 假设序列如下

$$x_t = \beta_0 + \beta_1 t + a_t$$

$$E(a_t) = 0, Var(a_t) = \sigma^2, Cov(a_t, a_{t-i}) = 0, \forall i \geq 1$$

- 考察一阶差分后序列和二阶差分序列的平稳性与方差

例5.4 差分比较

- 一阶差分
  - 平稳

$$\nabla x_t = x_t - x_{t-1}$$
$$= \beta_1 + a_t - a_{t-1}$$

  - 方差小

$$Var(\nabla x_t) = Var(a_t - a_{t-1})$$
$$= 2\sigma^2$$

- 二阶差分（过差分）
  - 平稳

$$\nabla^2 x_t = \nabla x_t - \nabla x_{t-1}$$
$$= a_t - 2a_{t-1} + a_{t-2}$$

  - 方差大

$$Var(\nabla^2 x_t) = Var(a_t - 2a_{t-1} + a_{t-2})$$
$$= 6\sigma^2$$

本章结构

1. Cramer分解
2. 差分运算
3. 单位根检验
4. ARIMA模型

单位根检验

- Dickdy&Fuller（1979）最早给出了一种单位根检验方法，这种方法现在被称为DF检验
- DF检验的构造思想是：假设序列 y_t 具有最简单的一阶线性差分方程结构

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

- 该方程的特征根为 $\lambda = \phi_1$ ，所以序列平稳性检验，就是检验 $|\phi_1|$ 是否小于1。上式等号两边同时减 $y_{t\text{-}1}$ ，该等式等价表达是

$$\nabla y_t = \rho y_{t-1} + \varepsilon_t$$

其中 $\nabla y_t = y_t - y_{t-1}$ ，数学上称为差分运算，定义 $\rho = |\phi_1| - 1$ 。

- 则序列平稳性检验等价于检验 ρ 值的大小：
  - 如果 $\rho \ge 0$，则序列非平稳。
  - 如果 $\rho < 0$，则序列平稳。

DF检验

- 假设条件

  原假设：序列非平稳H_0：$\rho \geq 0$

  备择假设：序列平稳 H_1：$\rho < 0$

- 检验统计量

$\rho = 0$ 时 $$t = \frac{\hat{\rho}}{S(\hat{\rho})} \xrightarrow{\text{渐近}} N（0,1）$$

$\rho \neq 0$ 时 $$\tau = \frac{\hat{\rho}}{S(\hat{\rho})} \xrightarrow{\text{极限}} DF\text{通过随机拟合得到的经验分布}$$

三种不同的序列产生机制

- DF检验只能把序列分为平稳或非平稳两大类
- 但其实序列真实的产生机制很复杂，Dickdy&Fuller（1979）考察了三种不同的回归机制产生的序列平稳性检验

$$(1)\, y_t = b y_{t-1} + \varepsilon_t$$

$$(2)\, y_t = a + b y_{t-1} + \varepsilon_t$$

$$(3)\, y_t = a + b y_{t-1} + ct + \varepsilon_t$$

DF检验的三种类型

- 第一种类型：无漂移项，无趋势（no drift， no trend）

$$\nabla x_t = \rho x_{t-1} + \varepsilon_t$$

- 第二种类型：有漂移项，无趋势（with drift， no trend）

$$\nabla x_t = a + \rho x_{t-1} + \varepsilon_t$$

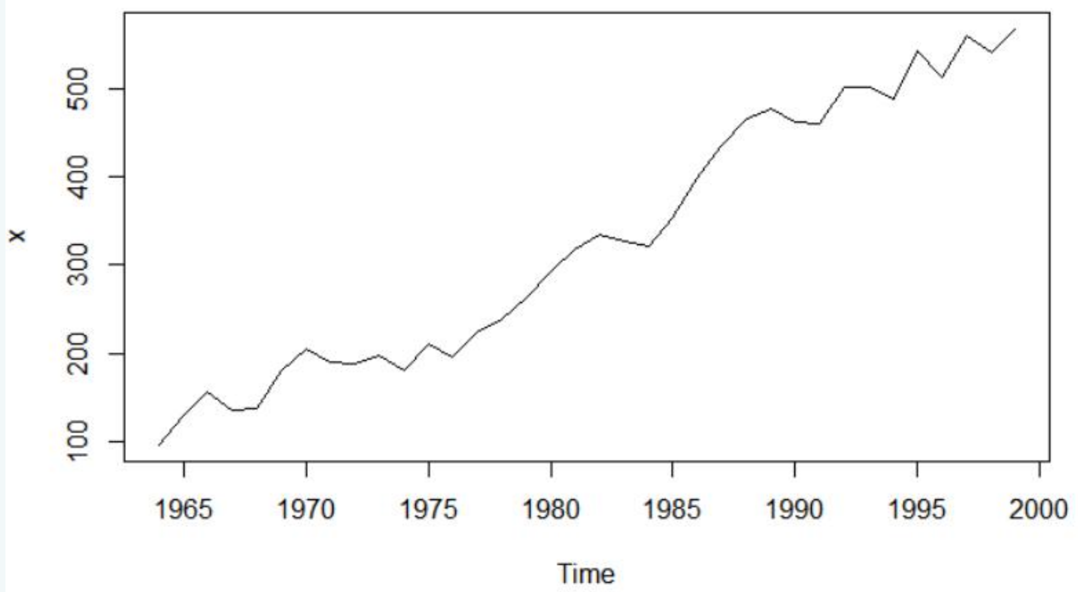
- 第三种类型：有常数，有趋势（with drift， with trend）

$$\nabla x_t = a + \rho x_{t-1} + ct + \varepsilon_t$$

- 这三种方程的检验原理相同，都是检验ρ是否为零，但是它们的临界值不同。

对中国纱产量序列的平稳性进行DF检验

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag  ADF p.value
[1,]   0 2.89     0.99
[2,]   1 3.06     0.99
Type 2: with drift no trend
     lag     ADF p.value
[1,]   0 -0.3453   0.905
[2,]   1 -0.0164   0.950
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -2.57   0.341
[2,]   1 -2.25   0.463
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```



对北京市每年最高气温的平稳性进行DF检验

Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend

	lag	ADF	p.value
[1,]	0	-0.3255	0.547
[2,]	1	-0.0289	0.632

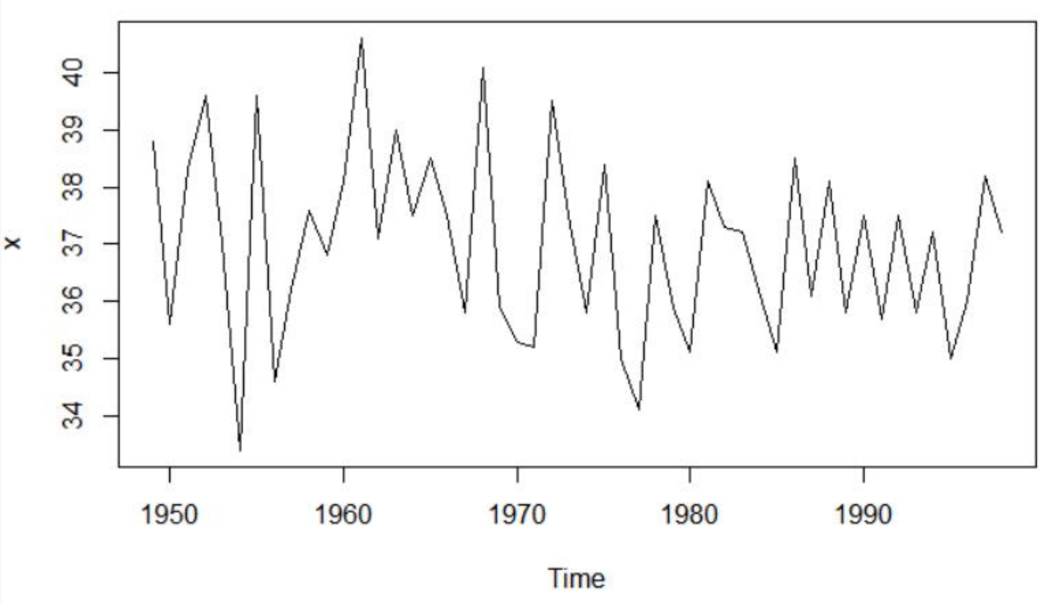
Type 2: with drift no trend

	lag	ADF	p.value
[1,]	0	-8.29	0.01
[2,]	1	-5.28	0.01

Type 3: with drift and trend

	lag	ADF	p.value
[1,]	0	-8.51	0.01
[2,]	1	-5.68	0.01

Note: in fact, p.value = 0.01 means p.value <= 0.01



ADF检验

- DF检验只适用于AR(1)过程的平稳性检验 。
- 为了使检验能适用于AR(p)过程的平稳性检验，人们对检验进行了一定的修正，得到增广检验（Augmented Dickey—Fuller），简记为ADF检验
- ADF检验以原理

$$
\begin{aligned}
y_t - y_{t-1} &= (\phi_1 - 1)y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} \\
&= (\phi_1 + \phi_2 + \cdots + \phi_p - 1)y_{t-1} - (\phi_2 + \cdots + \phi_p)\nabla y_{t-1} - (\phi_3 + \cdots + \phi_p)\nabla y_{t-2} + \cdots + \phi_{p-1}\nabla y_{t-p-1}
\end{aligned}
$$

- 记 $\rho = \phi_1 + \phi_2 + \cdots + \phi_p - 1$，序列是否平稳，等价于检验 ρ 的值

ADF检验

- 假设条件

  原假设：序列非平稳H_0： $\rho \geq 0$

  备择假设：序列平稳 H_1： $\rho < 0$

  其中： $\rho = \phi_1 + \phi_2 + \cdots + \phi_p - 1$

- 检验统计量

$$\tau = \frac{\hat{\rho}}{S(\hat{\rho})}$$

ADF检验的三种类型

- 第一种类型：无漂移项，无趋势（no drift， no trend）

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$$

- 第二种类型：有漂移项，无趋势（with drift， no trend）

$$x_t = \mu + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$$

- 第三种类型：有常数，有趋势（with drift， with trend）

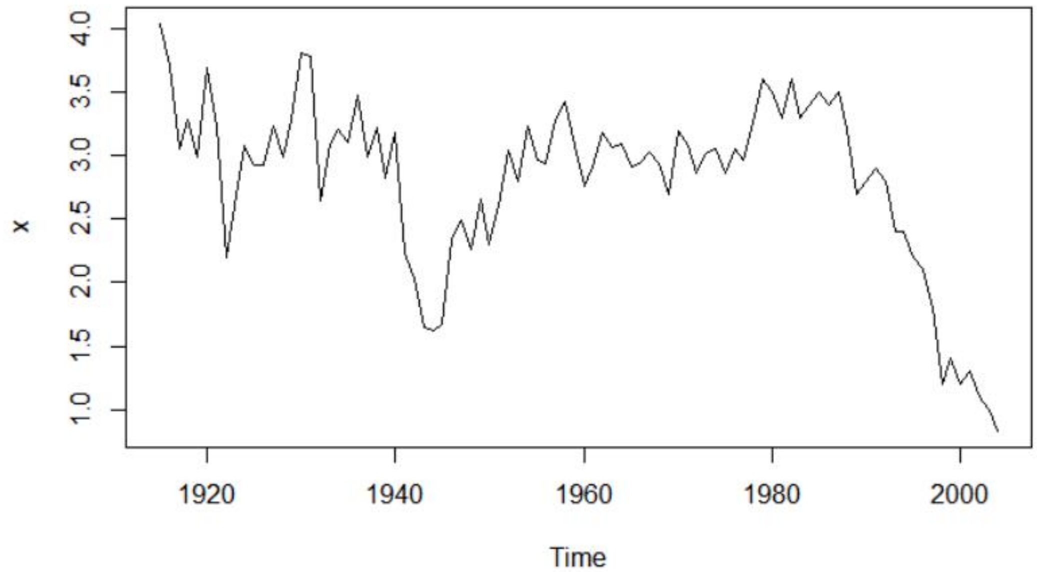
$$x_t = \mu + \beta t + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$$

对中国纱产量序列的平稳性进行ADF检验

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag  ADF p.value
[1,]   0 2.89    0.99
[2,]   1 3.06    0.99
[3,]   2 2.56    0.99
[4,]   3 2.34    0.99
Type 2: with drift no trend
     lag     ADF p.value
[1,]   0 -0.3453   0.905
[2,]   1 -0.0164   0.950
[3,]   2  0.2240   0.968
[4,]   3 -0.0663   0.943
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -2.57   0.341
[2,]   1 -2.25   0.463
[3,]   2 -2.57   0.341
[4,]   3 -2.04   0.542
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```

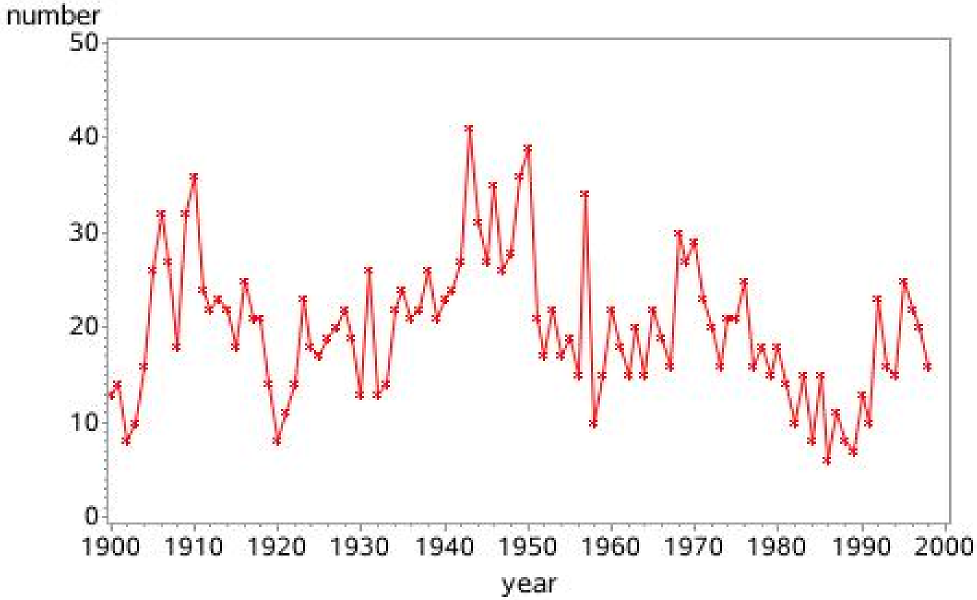
对澳大利亚死亡率序列的平稳性进行ADF检验

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -1.39   0.179
[2,]   1 -1.32   0.204
[3,]   2 -1.15   0.265
[4,]   3 -1.16   0.263
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -1.978   0.337
[2,]   1 -1.312   0.586
[3,]   2 -0.524   0.863
[4,]   3 -0.522   0.864
Type 3: with drift and trend
     lag    ADF p.value
[1,]   0 -2.294   0.449
[2,]   1 -1.645   0.719
[3,]   2 -0.951   0.941
[4,]   3 -0.906   0.948
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```



对全球7+地震序列的平稳性进行ADF检验

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag    ADF p.value
[1,]   0 -1.585   0.109
[2,]   1 -1.045   0.304
[3,]   2 -0.653   0.445
[4,]   3 -0.509   0.497
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -5.35  0.0100
[2,]   1 -3.92  0.0100
[3,]   2 -3.18  0.0245
[4,]   3 -2.96  0.0450
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -5.55  0.0100
[2,]   1 -4.14  0.0100
[3,]   2 -3.51  0.0448
[4,]   3 -3.37  0.0634
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```



本章结构

1. Cramer分解
2. 差分运算
3. 单位根检验
4. ARIMA模型

ARIMA模型结构

- 使用场合
  - 差分平稳序列拟合

- 模型结构

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \quad Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ Ex_s \varepsilon_t = 0, \forall s < t \end{cases}$$

随机游走模型(random walk)

- 模型结构

$$\begin{cases} x_t = x_{t-1} + \varepsilon_t \\ E(\varepsilon_t) = 0, \quad Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E x_s \varepsilon_t = 0, \forall s < t \end{cases}$$

- 模型使用场合
  - Karl Pearson(1905)在《自然》杂志上提问：假如有个醉汉醉得非常严重，完全丧失方向感，把他放在荒郊野外，一段时间之后再去找他，在什么地方找到他的概率最大呢？这个醉汉的行走轨迹就是一个随机游走模型。
  - 传统的经济学家普遍认为投机价格的走势类似于随机游走模型，随机游走模型也是有效市场理论的核心。

ARIMA模型的平稳性

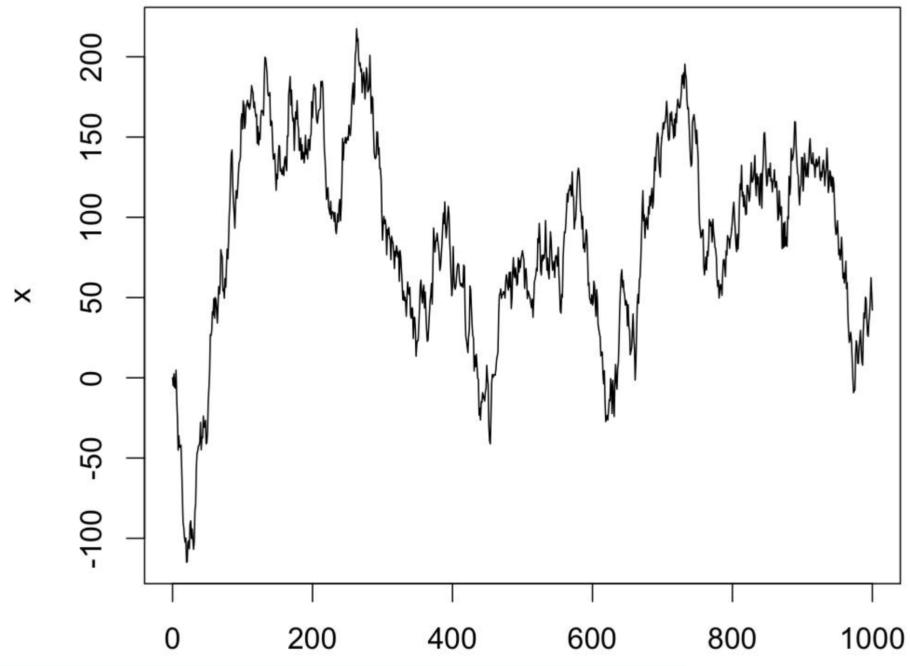
- ARIMA(p,d,q)模型共有p+d个特征根，其中p个在单位圆内，d个在单位圆上。所以当

$$d \neq 0$$

时ARIMA(p,d,q)模型非平稳。

- 例4.8

ARIMA(0,1,0)时序图



ARIMA模型的方差齐性

- $d \neq 0$ 时，原序列方差非齐性

$$ARIMA(0,1,0)\text{模型}$$

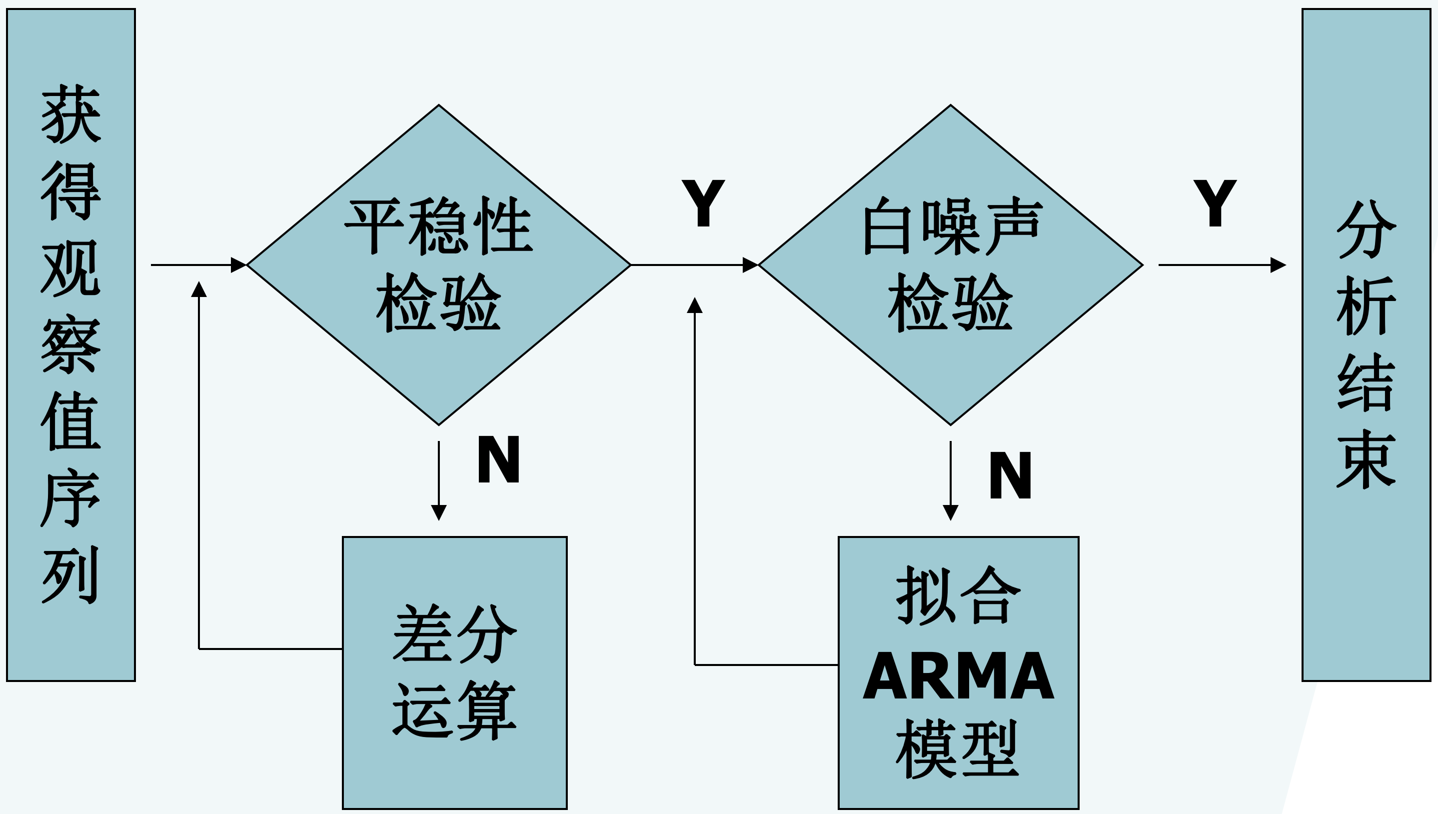
$$Var(x_t) = Var(x_0 + \varepsilon_t + \varepsilon_{t-1} + \cdots \varepsilon_1) = t\sigma_\varepsilon^2$$

- d阶差分后，差分后序列方差齐性

$$ARIMA(0,1,0)\text{模型}$$

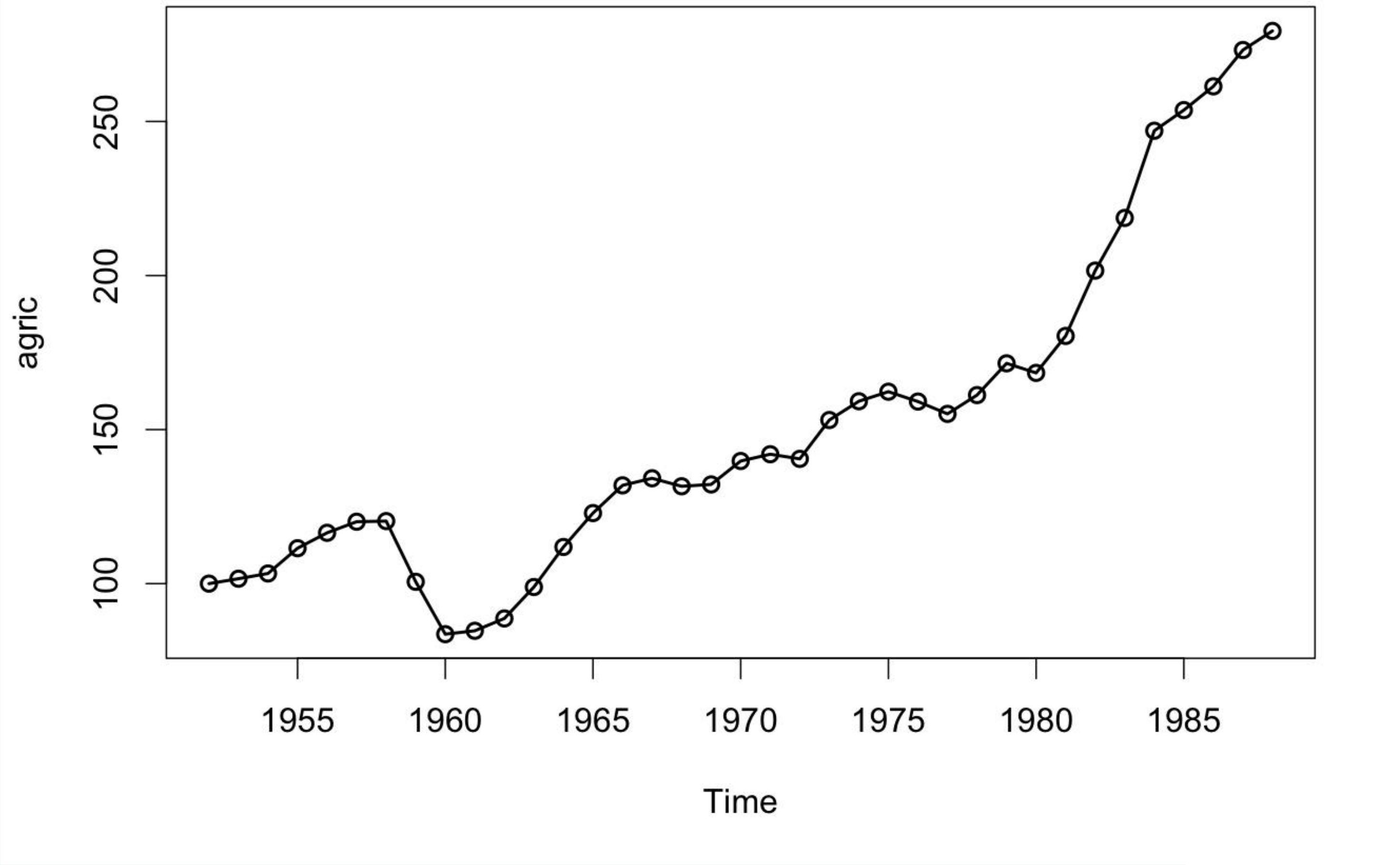
$$Var(\nabla x_t) = Var(\varepsilon_t) = \sigma_\varepsilon^2$$

ARIMA模型建模步骤

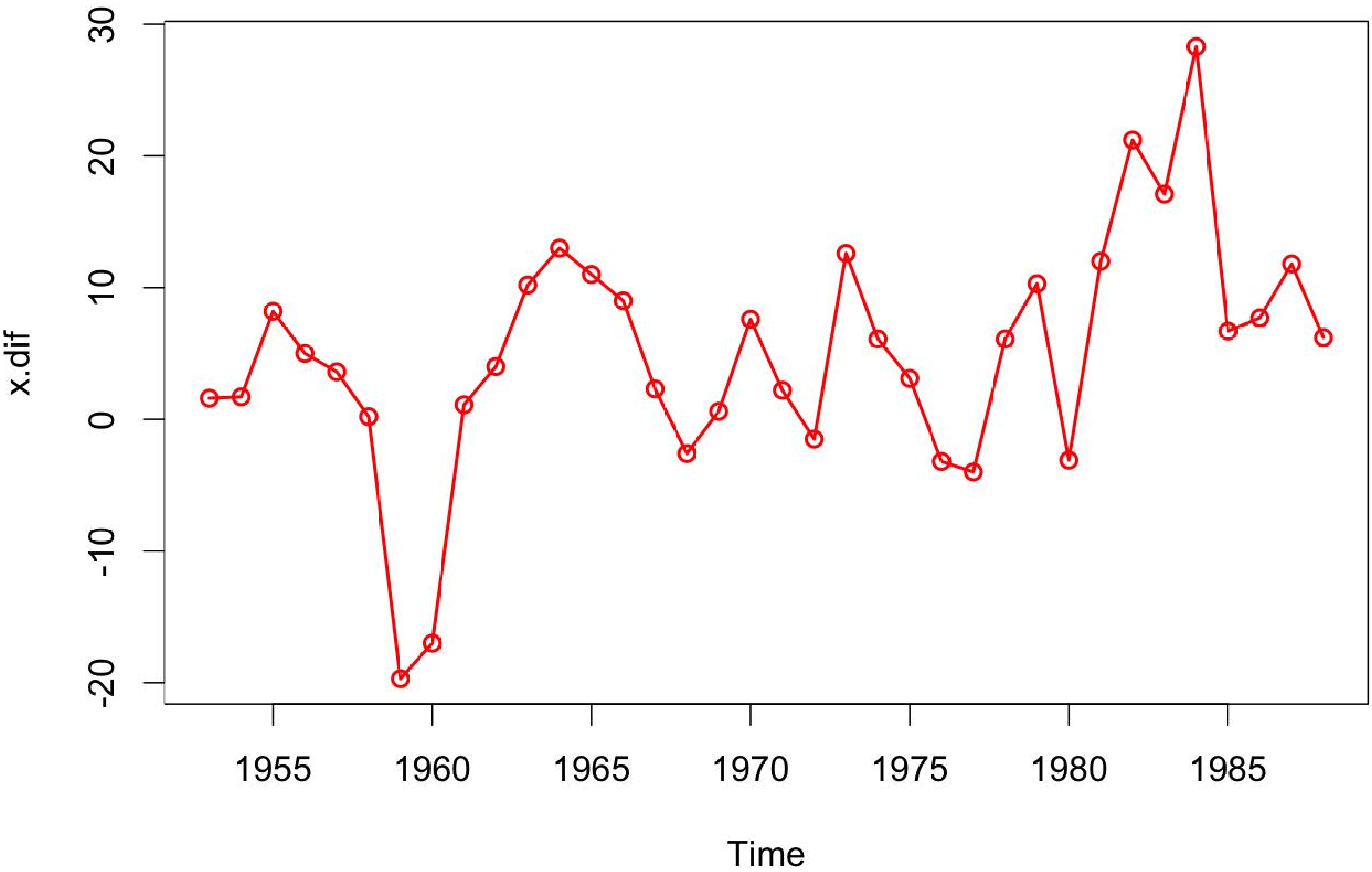


例5.9

- 对1952年——1988年中国农业实际国民收入指数序列建模



一阶差分序列时序图



一阶差分序列ADF检验

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -2.65  0.0100
[2,]   1 -2.42  0.0185
[3,]   2 -1.96  0.0491
[4,]   3 -2.17  0.0316
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -3.16  0.0353
[2,]   1 -3.01  0.0469
[3,]   2 -2.44  0.1666
[4,]   3 -2.85  0.0672
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -3.52  0.0559
[2,]   1 -3.49  0.0601
[3,]   2 -3.18  0.1127
[4,]   3 -4.16  0.0147
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```

一阶差分后序列白噪声检验

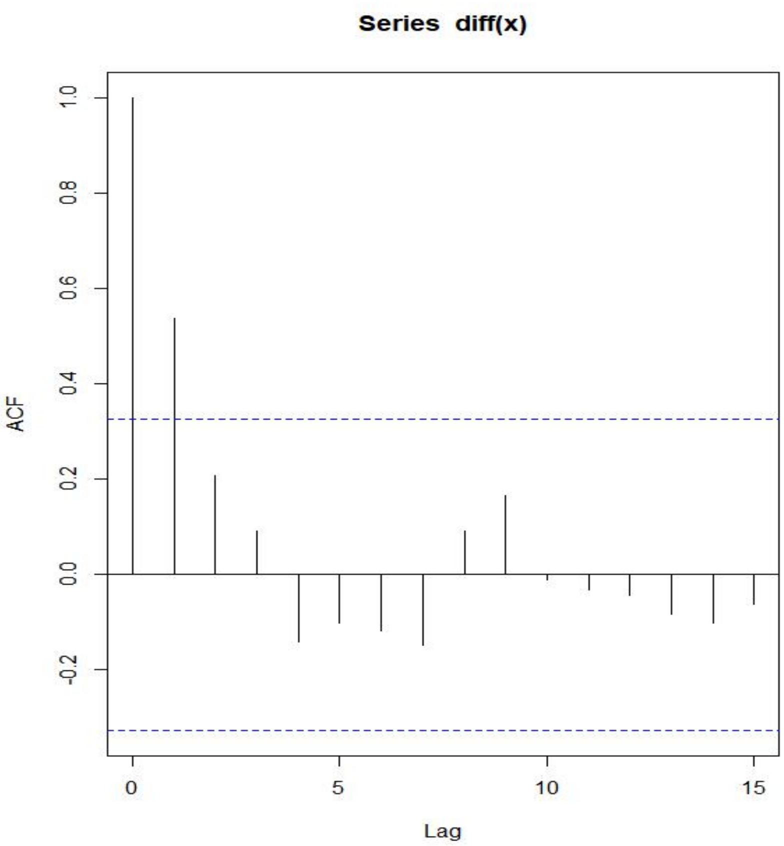
Box-Pierce test

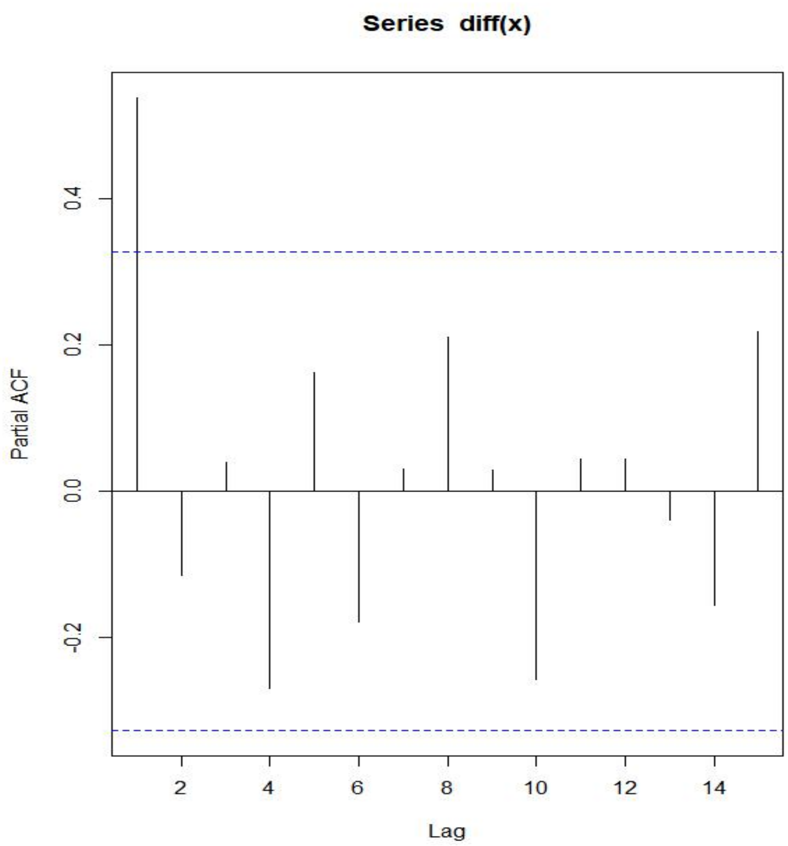
data: diff(x)

X-squared = 12.257, df = 3, p-value = 0.006553

拟合ARMA模型

❖ 自相关图与偏自相关图





建模

- 定阶
  - ARIMA(1,1,0)或者ARIMA(0,1,1)
- 参数估计

  model1： $\nabla x_t = 0.6353 x_{t-1} + \varepsilon_t$

  model2： $\nabla x_t = \varepsilon_t + 0.7355 \varepsilon_{t-1}$
- 模型检验
  - 两个模型均显著成立，参数均显著非零
- 模型优化
  - AIC和BIC检验，model1要偏小一点
- 序列预测

ARIMA模型预测

- 原则
  - 最小均方误差预测原理

- Green函数递推公式

$$
\begin{cases}
\psi_1 = \phi_1 - \theta_1 \\
\psi_2 = \phi_1 \psi_1 + \phi_2 - \theta_2 \\
\vdots \\
\psi_j = \phi_1 \psi_{j-1} + \cdots + \phi_{p+d} \psi_{j-p-d} - \theta_j
\end{cases}
$$

预测值

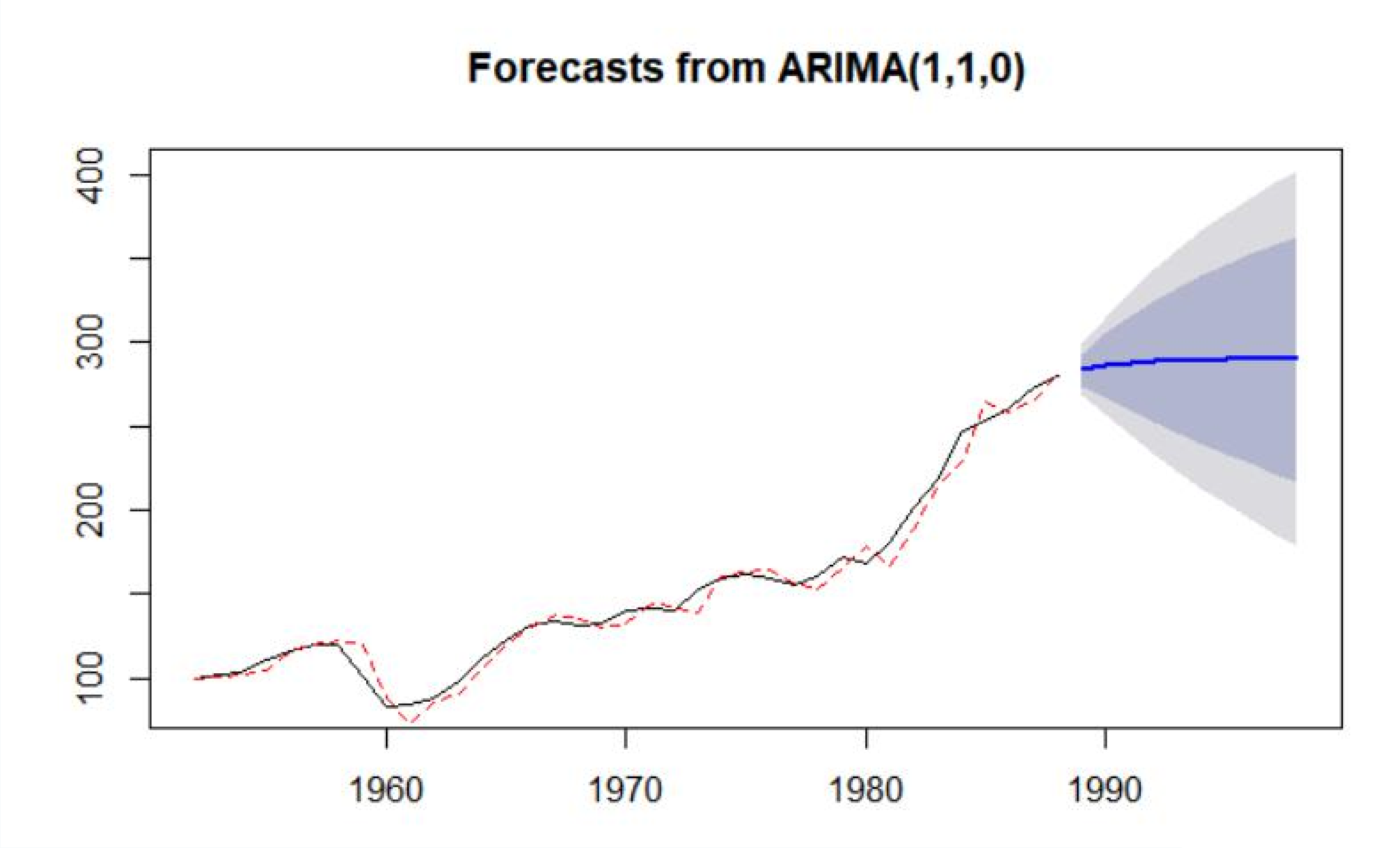
$$x_{t+l} = (\varepsilon_{t+l} + \psi_1\varepsilon_{t+l-1} + \cdots + \psi_{l-1}\varepsilon_{t+1}) + (\psi_l\varepsilon_t + \psi_{l+1}\varepsilon_{t-1} + \cdots)$$

$$\underbrace{\varepsilon_{t+l} + \psi_1\varepsilon_{t+l-1} + \cdots + \psi_{l-1}\varepsilon_{t+1}}_{e_t(l)} \qquad \underbrace{\psi_l\varepsilon_t + \psi_{l+1}\varepsilon_{t-1} + \cdots}_{\hat{x}_t(l)}$$

$$E[e_t(l)] = 0$$

$$Var[e_t(l)] = (1 + \psi_1^2 + \cdots + \psi_{l-1}^2)\sigma_\varepsilon^2$$

例6.9续：对中国农业实际国民收入指数序列的预测



谢谢！