

# 无季节效应的非平稳序列分析

---

05

# 本章内容

01

**Cramer分解定理**

02

**差分平稳**

03

**ARIMA模型**

04

**疏系数模型**

# Cramer分解定理

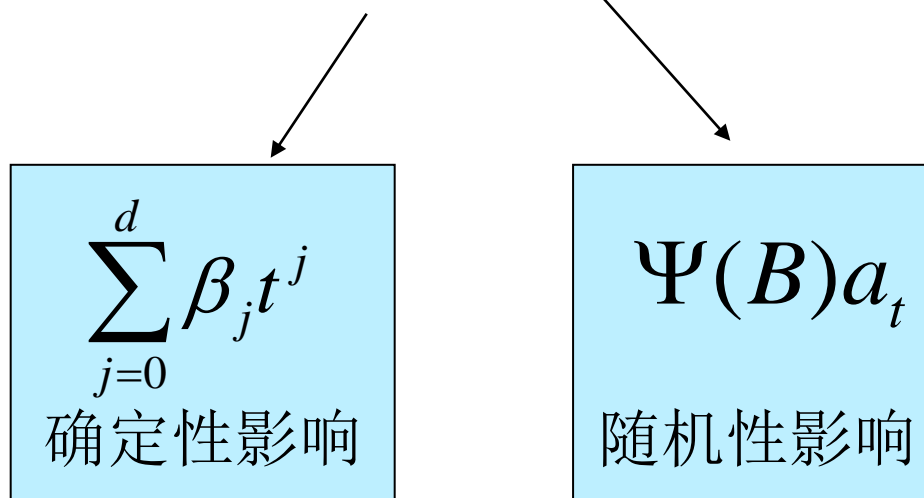
- Cramer分解定理
  - Harald Cramer(1893-1985)。瑞典人，斯德哥尔摩大学教授，著名的统计学家和保险精算学家。
  - Cramer 分解定理是Wold分解定理的推广。Wold分解定理是平稳序列的理论基础，Cramer分解定理是非平稳序列的理论基础。
  - Cramer 分解定理说明任何一个序列的波动都可以视为同时受到了确定性影响和随机性影响的综合作用。平稳序列要求这两方面的影响都是稳定的，而非平稳序列产生的机理就在于它所受到的这两方面的影响至少有一方面是不稳定的。



# Cramer分解定理

- 任何一个时间序列  $\{x_t\}$  都可以分解为两部分的叠加：
  - 一部分是由时间  $t$  的多项式决定的确定性成分
  - 另一部分是由白噪声序列决定的随机性成分

$$x_t = \mu_t + \varepsilon_t$$



# 本章内容

01

**Cramer分解定理**

02

**差分平稳**

03

**ARIMA模型**

04

**疏系数模型**

# 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

$$\nabla^d \sum_{j=0}^d \beta_j t^j = c, \quad c \text{ 为某一常数}$$

- 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

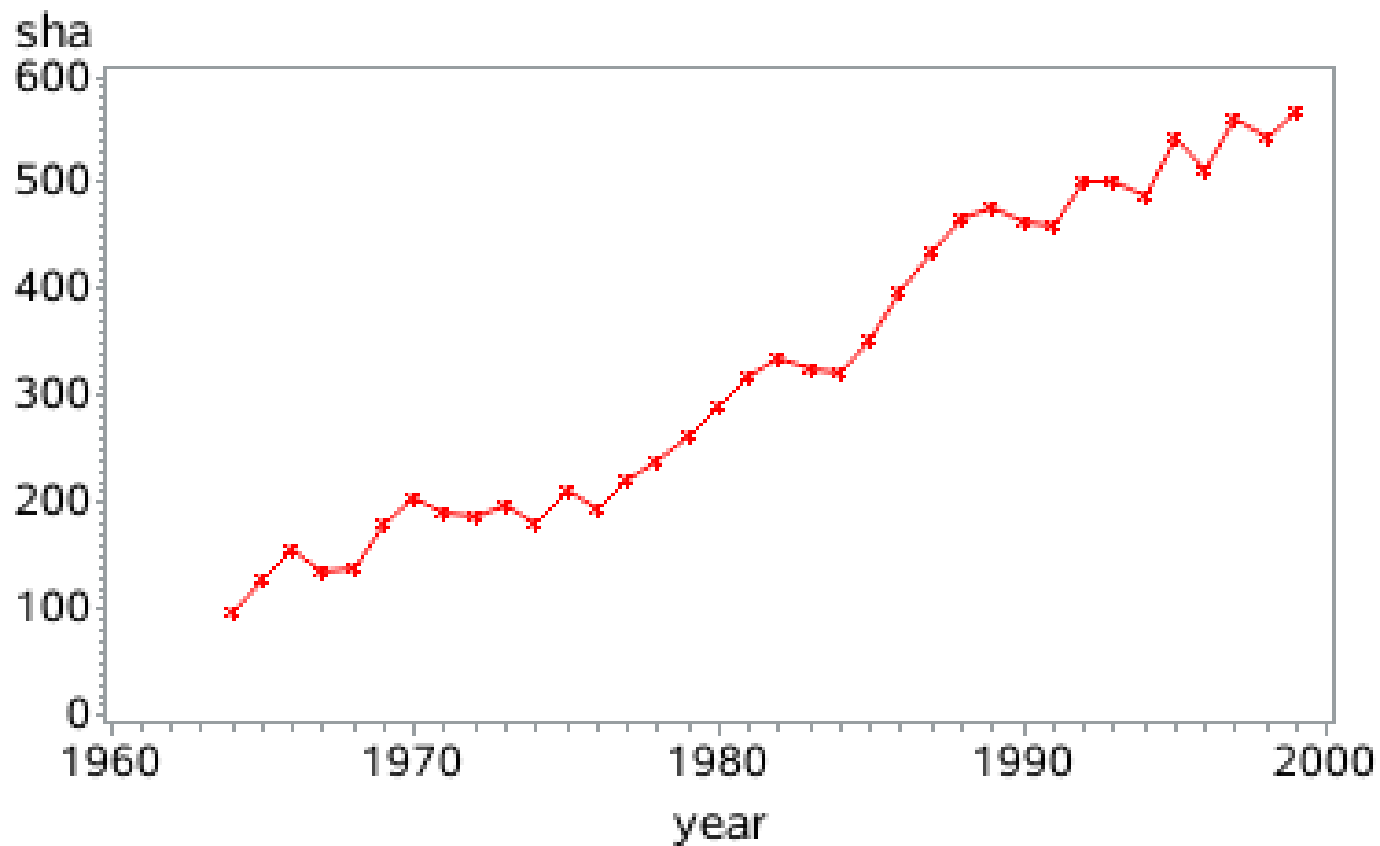
---

- 序列蕴含着显著的线性趋势，通常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

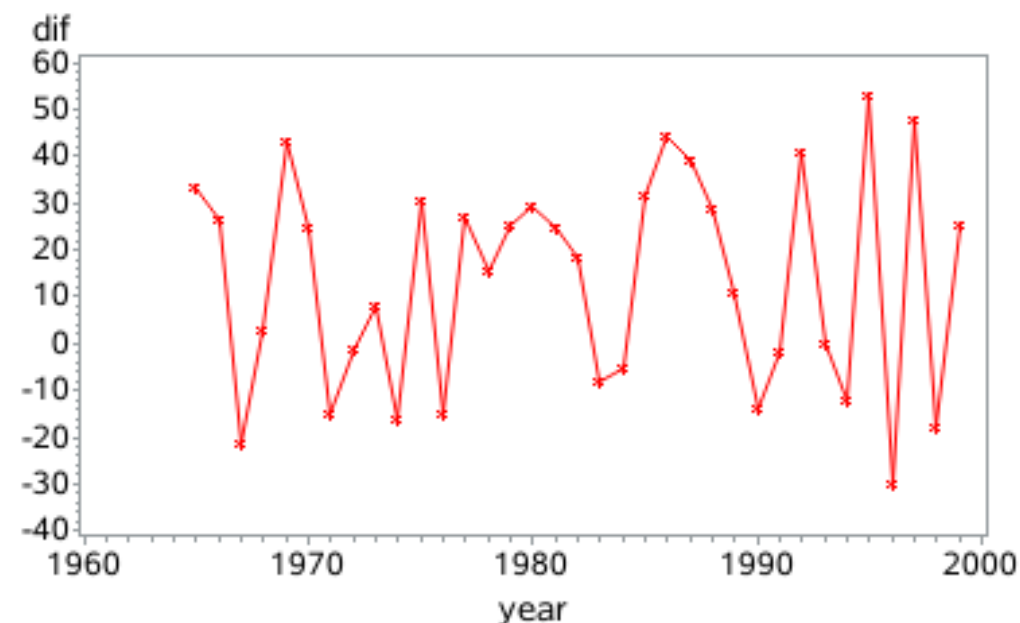
- 尝试提取1964-1999年中国纱年产量序列中的确定性信息。



## 例5-1 一阶差分提取确定性信息

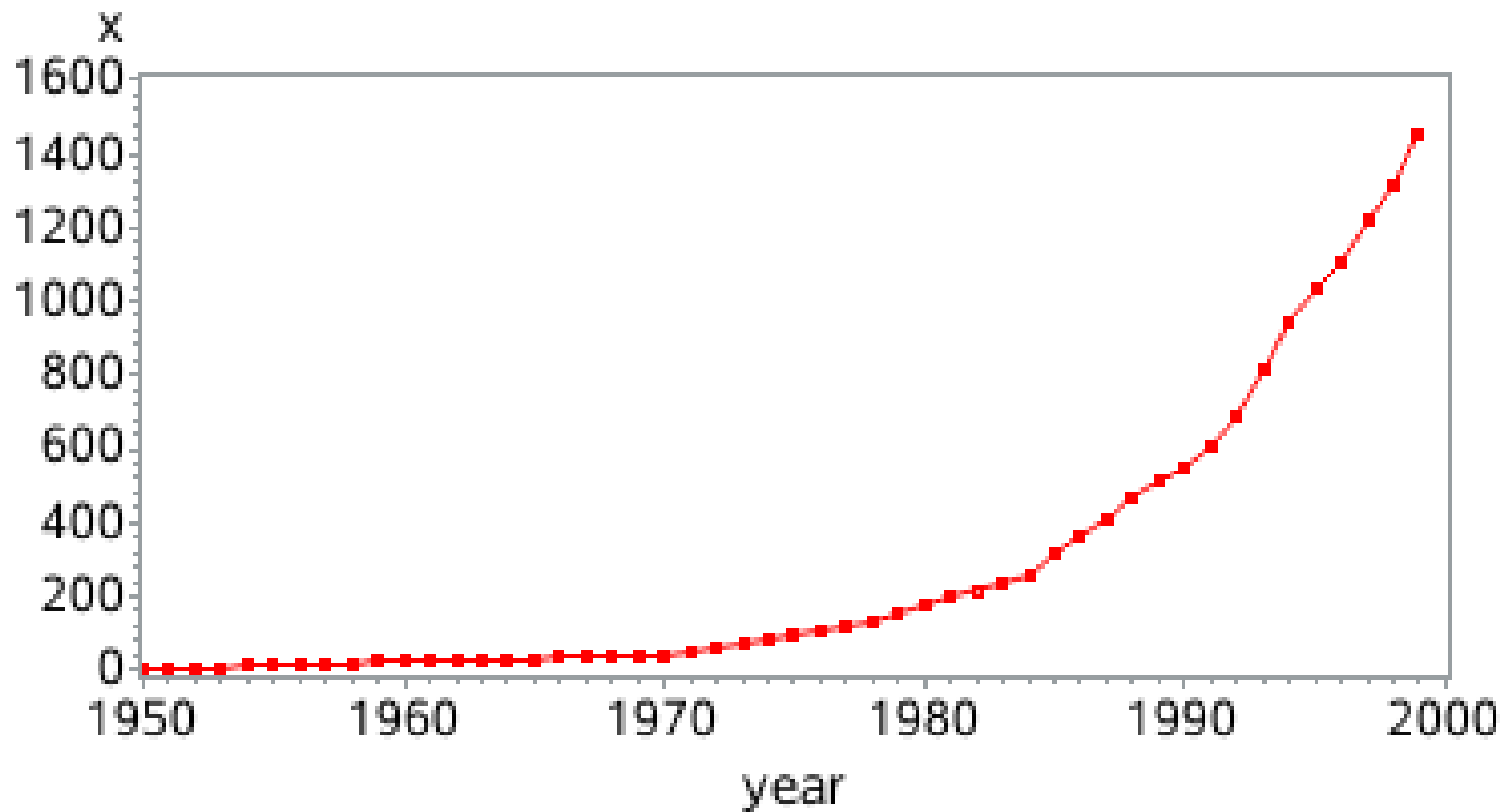
- 中国纱产量序列蕴涵着显著的线性递增趋势。对该序列进行 1 阶差分提取线性趋势信息。1 阶差分后时序图如右图所示。
- 差分后时序图显示：1 阶差分后序列呈现出非常平稳的波动特征，这说明 1 阶差分运算非常成功地从原序列中提取出线性趋势。

一阶差分后时序图



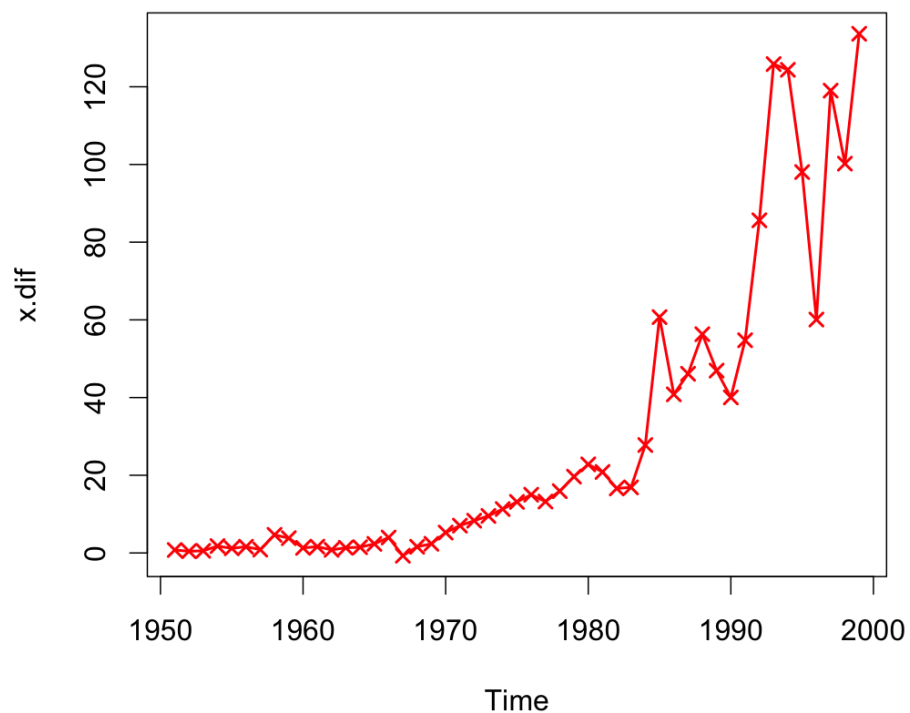
## 例5-2

- 尝试提取1950—1999年北京市民用车辆拥有量序列的确定性信息



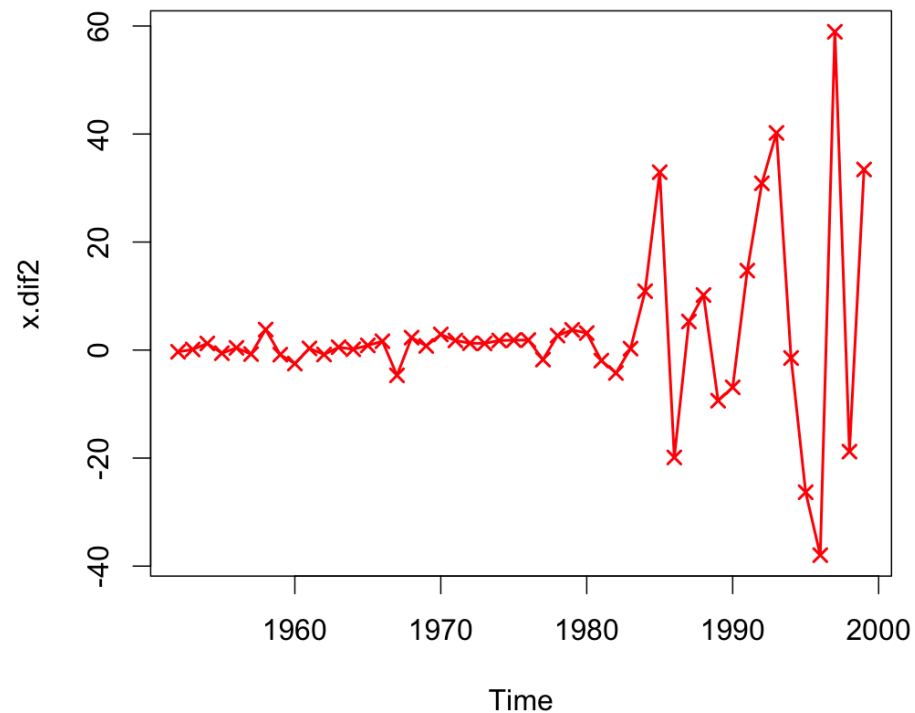
# 差分后序列时序图

- 一阶差分



1阶差分没有实现趋势平稳

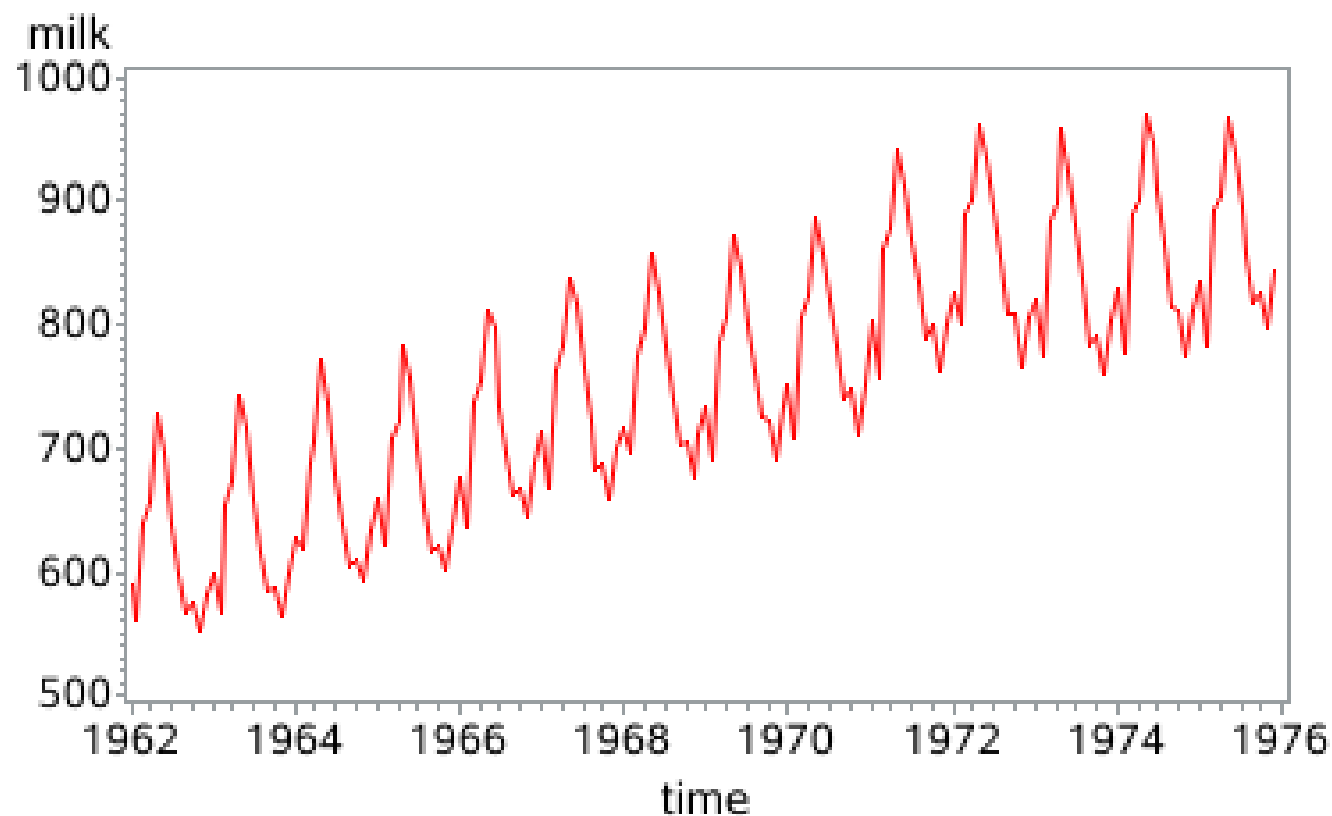
- 二阶差分



2阶差分实现趋势平稳，但随机波动不平稳

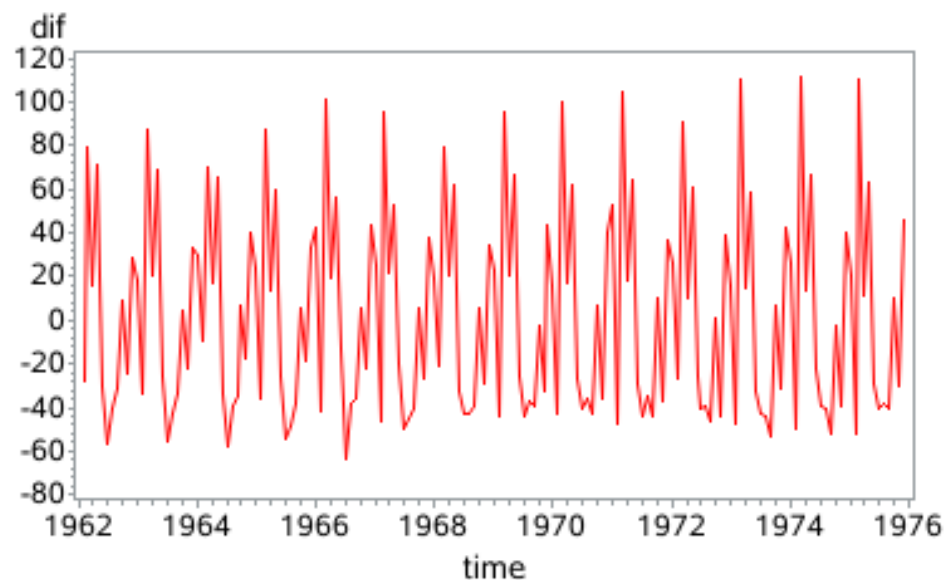
## 例5-3

- 利用差分运算提取1962年1月至1975年12月平均每头奶牛的月产奶量序列中的确定性信息。



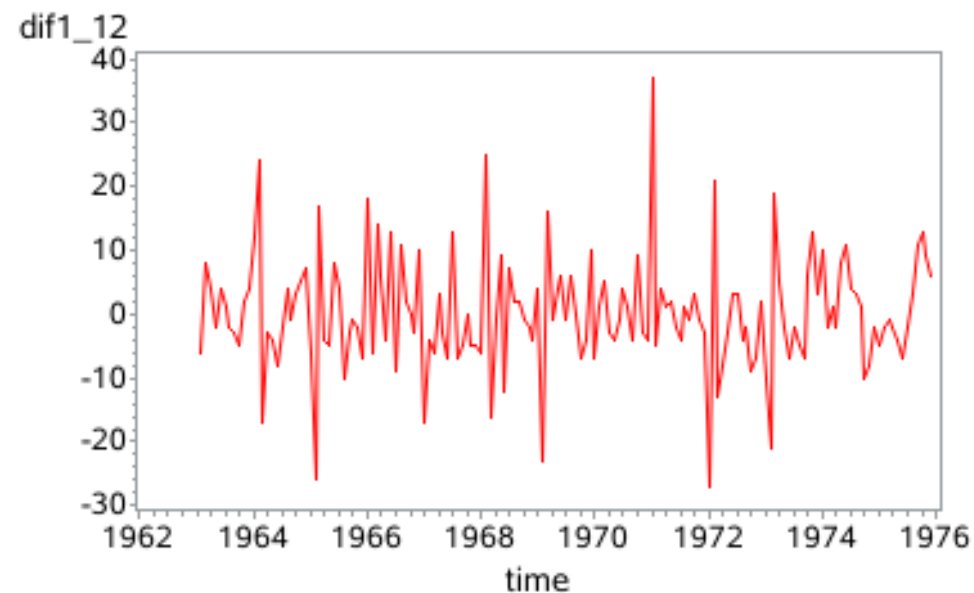
# 差分后序列时序图

- 一阶差分



1 阶差分后线性递增信息被提取，  
残留稳定的季节波动和随机波动。

- 1 阶 - 12 步差分



周期差分可以非常好地提取周  
期信息

# 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

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

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

- 考察一阶差分后序列和二阶差分序列的平稳性与方差



# 比较

- 一阶差分

- 平稳

$$\begin{aligned}\nabla x_t &= x_t - x_{t-1} \\ &= \beta_1 + a_t - a_{t-1}\end{aligned}$$

- 方差小

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

- 二阶差分（过差分）

- 平稳

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

- 方差大

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

# 本章内容

01

**Cramer分解定理**

02

**差分平稳**

03

**ARIMA模型**

04

**疏系数模型**

# ARIMA模型结构

---

- 使用场合
  - 差分平稳序列的拟合
- 模型结构

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

## 特殊的ARIMA模型

---

❖  $d=0$

$$\text{ARIMA}(p,0,q)=\text{ARMA}(p,q)$$

❖  $p=0$

$$\text{ARIMA}(0,d,q)=\text{IMA}(d,q)$$

❖  $q=0$

$$\text{ARIMA}(p,d,0)=\text{ARI}(p,d)$$

❖  $d=1, P=q=0$

$$\text{ARIMA}(p,d,q)=\text{random walk model (随机游走模型)}$$

# 随机游走模型

- 模型结构

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

- 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

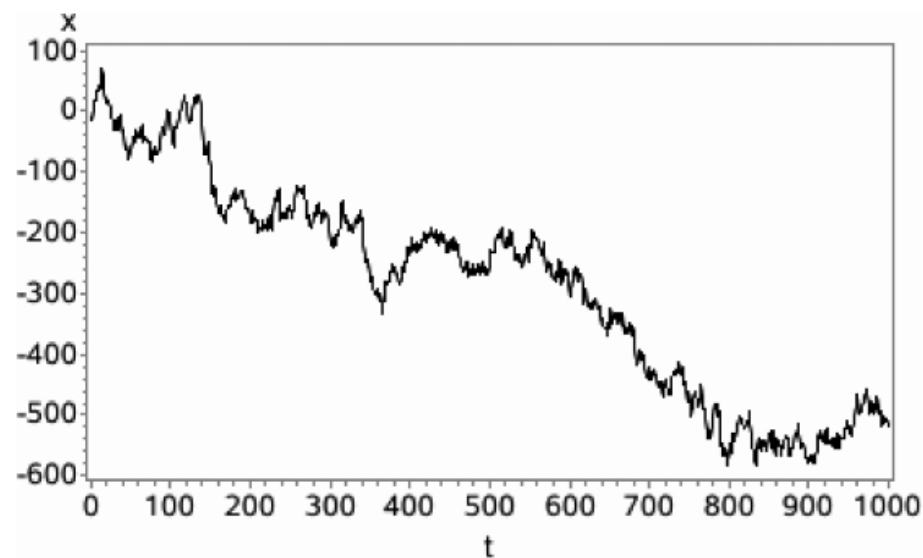
- ARIMA( $p, d, q$ )模型共有 $p + d$ 个特征根，其中 $p$ 个在单位圆内， $d$ 个在单位圆上。所以当

$$d \neq 0$$

时ARIMA( $p, d, q$ )模型非平稳。

- 例4.8

ARIMA(0,1,0)拟合序列时序图



# ARIMA模型的方差齐性 (以ARIMA(0,1,0) 为例)

- $d \neq 0$  时, 原序列方差非齐性

$$\begin{aligned}x_t &= x_{t-1} + \epsilon_t \\&= x_{t-2} + \epsilon_t + \epsilon_{t-1} \\&\vdots \\&= x_0 + \epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_1\end{aligned}$$

$$\begin{aligned}\text{Var}(x_t) &= \text{Var}(x_0 + \epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_1) \\&= t\sigma_\epsilon^2\end{aligned}$$

- $d$ 阶差分后, 差分后序列方差齐性

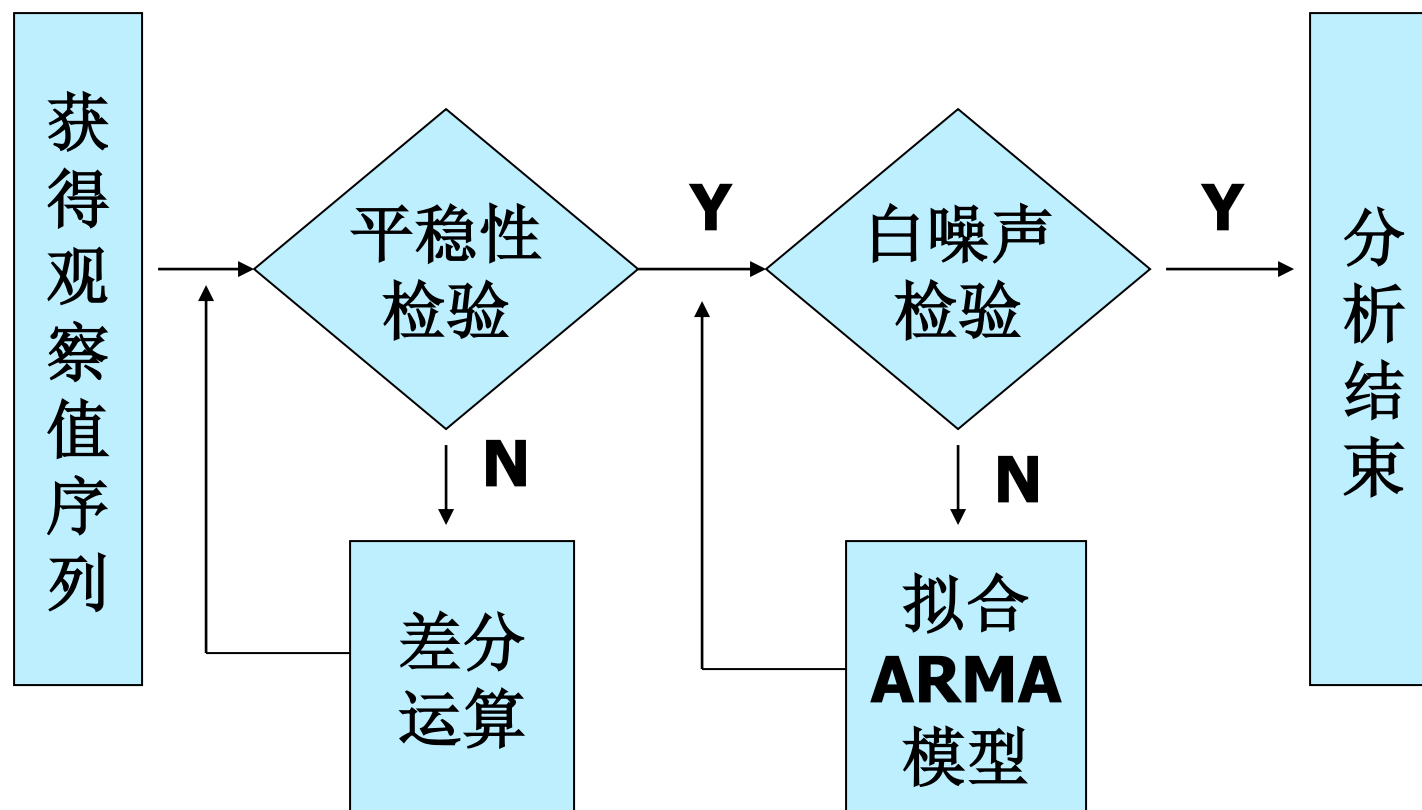
$$\nabla x_t = \epsilon_t$$

差分后序列方差齐性

$$\text{Var}(\nabla x_t) = \sigma_\epsilon^2$$



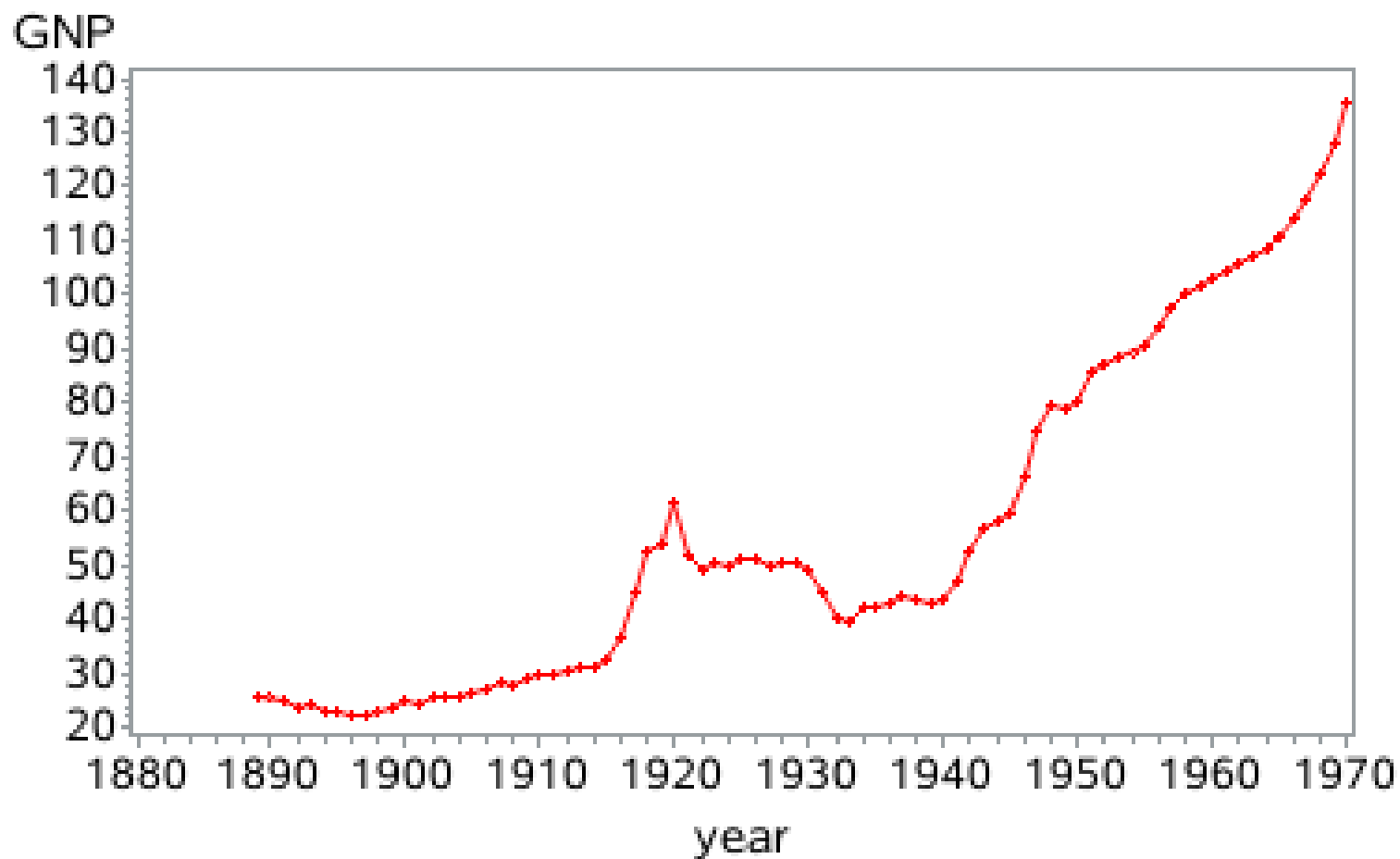
# ARIMA模型建模步骤





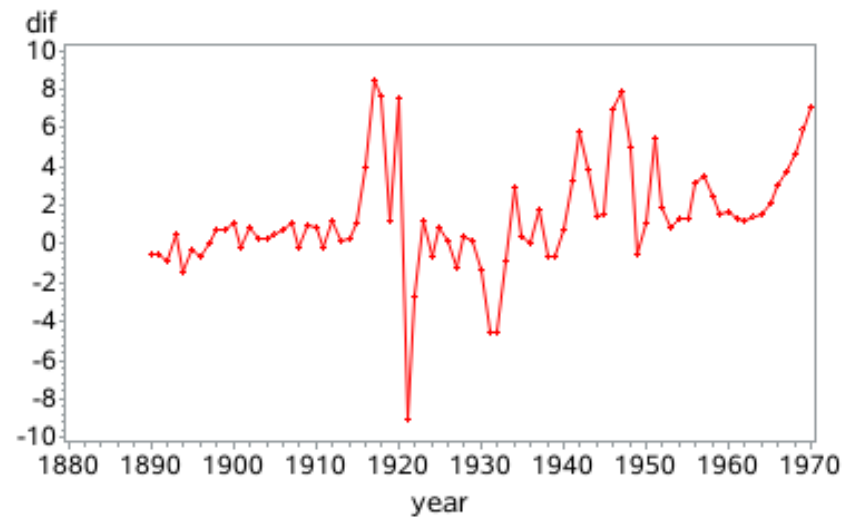
## 例5-6

- 对1889—1970年美国国民生产总值平减指数序列建模



# 差分平稳

- 一阶差分时序图



结论：  
平稳非白噪声序列

- ADF 检验

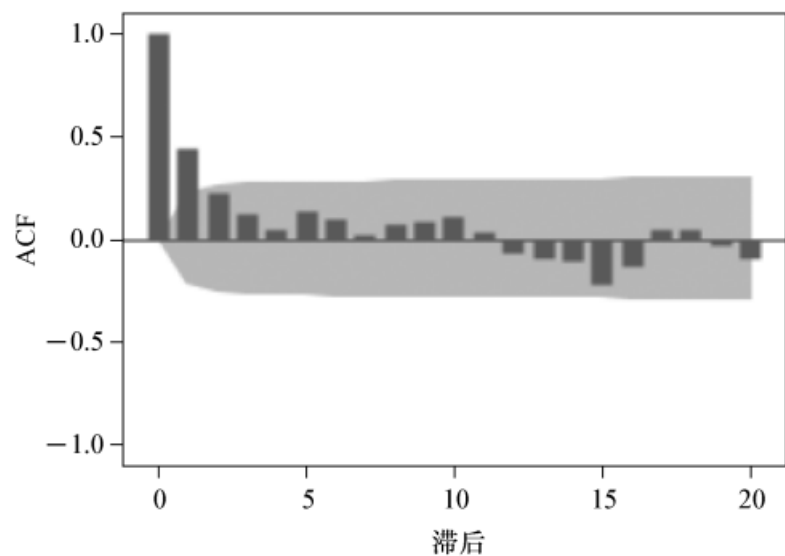
类型	延迟阶数	$\tau$ 统计量的值	$Pr < \tau$
类型一	0	-4.38	<0.000 1
	1	-3.25	0.001 5
	2	-2.61	0.009 7
类型二	0	-5.14	0.000 1
	1	-4.03	0.002 1
	2	-3.44	0.012 4
类型三	0	-5.73	<0.000 1
	1	-4.62	0.001 9
	2	-4.03	0.011 4

- 白噪声检验

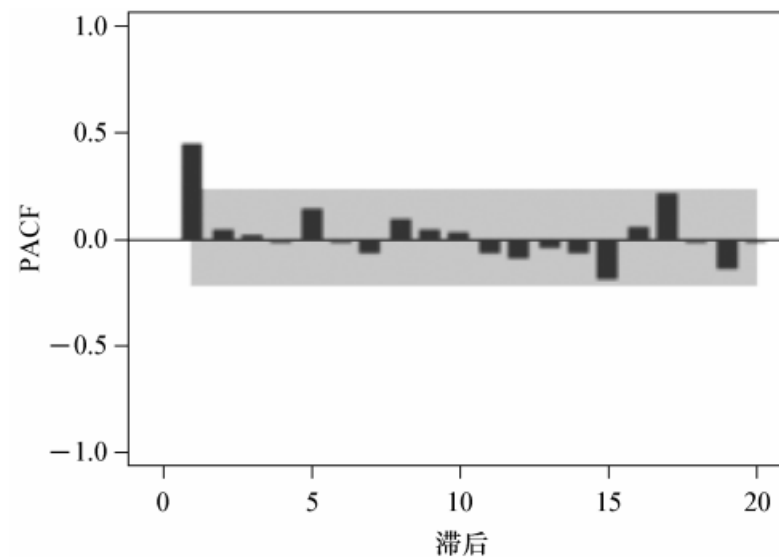
延迟阶数	纯随机性检验	
	LB 检验统计量的值	P 值
6	25.33	0.000 3
12	28.09	0.005 4
18	37.18	0.005 0

# 模型定阶

- 一阶差分后序列自相关图



- 一阶差分后序列偏自相关图



1 阶差分后序列的自相关图显示拖尾特征，偏自相关图显示 1 阶截尾特征。所以考虑用  $AR(1)$  模型拟合 1 阶差分后序列。考虑到前面已经进行的 1 阶差分运算，实际上使用  $ARIMA(1,1,0)$  模型拟合原序列。

# 模型拟合

---

- 基于条件最小二乘估计法得到的模型口径

$$\nabla x_t = 1.3746 + \frac{\epsilon_t}{1 - 0.4695B}$$

或等价表示为：

$$x_t = 0.7292 + 1.4695x_{t-1} - 0.4695x_{t-2} + \epsilon_t$$

式中，  $\text{Var}(\epsilon_t) = 6.3564$

# 模型检验

- 模型显著性检验

残差白噪声检验		
延迟阶数	LB 统计量	$P$ 值
6	2.68	0.748 9
12	4.93	0.934 4

- 参数显著性检验

参数显著性检验		
参数	$t$ 统计量	$P$ 值
$\mu$	2.64	0.009 9
$\phi_1$	4.57	$<0.000\ 1$

# ARIMA模型预测

- ARIMA (p,d,q)模型的一般表示方法为  $\Phi(B)(1-B)^d x_t = \Theta(B)\epsilon_t$   
如果把  $\Phi^*(B)$  记为广义自相关函数, 有

$$\Phi^*(B) = \Phi(B)(1-B)^d = 1 - \tilde{\phi}_1 B - \tilde{\phi}_2 B^2 - \dots$$

- 和ARMA模型一样, 也可以用随机扰动项的线性函数表示它 (Green函数)

$$x_t = \epsilon_t + \Psi_1 \epsilon_{t-1} + \Psi_2 \epsilon_{t-2} + \dots$$

- ARIMA模型Green函数递推公式

$$\begin{cases} \Psi_1 = \tilde{\phi}_1 - \theta_1 \\ \Psi_2 = \tilde{\phi}_1 \Psi_1 + \tilde{\phi}_2 - \theta_2 \\ \vdots \\ \Psi_j = \tilde{\phi}_1 \Psi_{j-1} + \dots + \tilde{\phi}_{p+d} \Psi_{j-p-d} - \theta_j \end{cases} \quad \text{式中,} \quad \Psi_j = \begin{cases} 0, & j < 0 \\ 1, & j = 0 \end{cases}, \quad \theta_j = 0, \quad j > q$$

# 预测值

$$x_{t+l} = (\epsilon_{t+l} + \Psi_1 \epsilon_{t+l-1} + \cdots + \Psi_{l-1} \epsilon_{t+1}) + (\Psi_l \epsilon_t + \Psi_{l+1} \epsilon_{t-1} + \cdots)$$



$$e_t(l)$$

$$e_t(l) = \epsilon_{t+l} + \Psi_1 \epsilon_{t+l-1} + \cdots + \Psi_{l-1} \epsilon_{t+1}$$

$$\text{Var}[e_t(l)] = (1 + \Psi_1^2 + \cdots + \Psi_{l-1}^2) \sigma_\epsilon^2$$



$$\hat{x}_t(l)$$

$$\hat{x}_t(l) = \Psi_l \epsilon_t + \Psi_{l+1} \epsilon_{t-1} + \Psi_{l+2} \epsilon_{t-2} + \cdots$$

## 例5-7

- 已知ARIMA (1,1,1)模型为  $(1-0.8B)(1-B)x_t = (1-0.6B)\epsilon_t$
- 已知:  $x_{t-1} = 4.5$   $x_t = 5.3$ ,  $\epsilon_t = 0.8$ ,  $\sigma_\epsilon^2 = 1$
- 求  $x_{t+3}$  的95%置信区间。

【解】

- 展开原模型, 得到等价模型  $x_t = 1.8x_{t-1} - 0.8x_{t-2} + \epsilon_t - 0.6\epsilon_{t-1}$
- 预测值的递推公式为

$$\hat{x}_t(1) = 1.8x_t - 0.8x_{t-1} - 0.6\epsilon_t = 5.46$$

$$\hat{x}_t(2) = 1.8\hat{x}_t(1) - 0.8x_t = 5.59$$

$$\hat{x}_t(3) = 1.8\hat{x}_t(2) - 0.8\hat{x}_t(1) = 5.69$$



## 例5-7解

- 3 期预测误差的方差为

$$\text{Var}[e(3)] = (1 + \Psi_1^2 + \Psi_2^2) \sigma_\varepsilon^2$$

- 广义自相关函数为

$$\begin{aligned}\Phi^*(B) &= \Phi(B)(1-B)^d \\ &= (1-0.8B)(1-B) \\ &= 1-1.8B+0.8B^2\end{aligned}$$

$$\text{则 } \tilde{\phi}_1 = 1.8, \tilde{\phi}_2 = -0.8$$

- Green函数为

$$\begin{cases} \Psi_1 = 1.8 - 0.6 = 1.2 \\ \Psi_2 = 1.8\Psi_1 - 0.8 = 1.36 \end{cases}$$

- 3期预测值方差为

$$\text{Var}[e(3)] = (1 + \Psi_1^2 + \Psi_2^2) \sigma_\varepsilon^2 = 4.2896$$

- 3期预测值的95%置信区间为

$$(\hat{x}_t(3) - 1.96 \sqrt{\text{Var}[e(3)]}, \hat{x}_t(3) + 1.96 \sqrt{\text{Var}[e(3)]}), \text{ 即 } (1.63, 9.75)$$

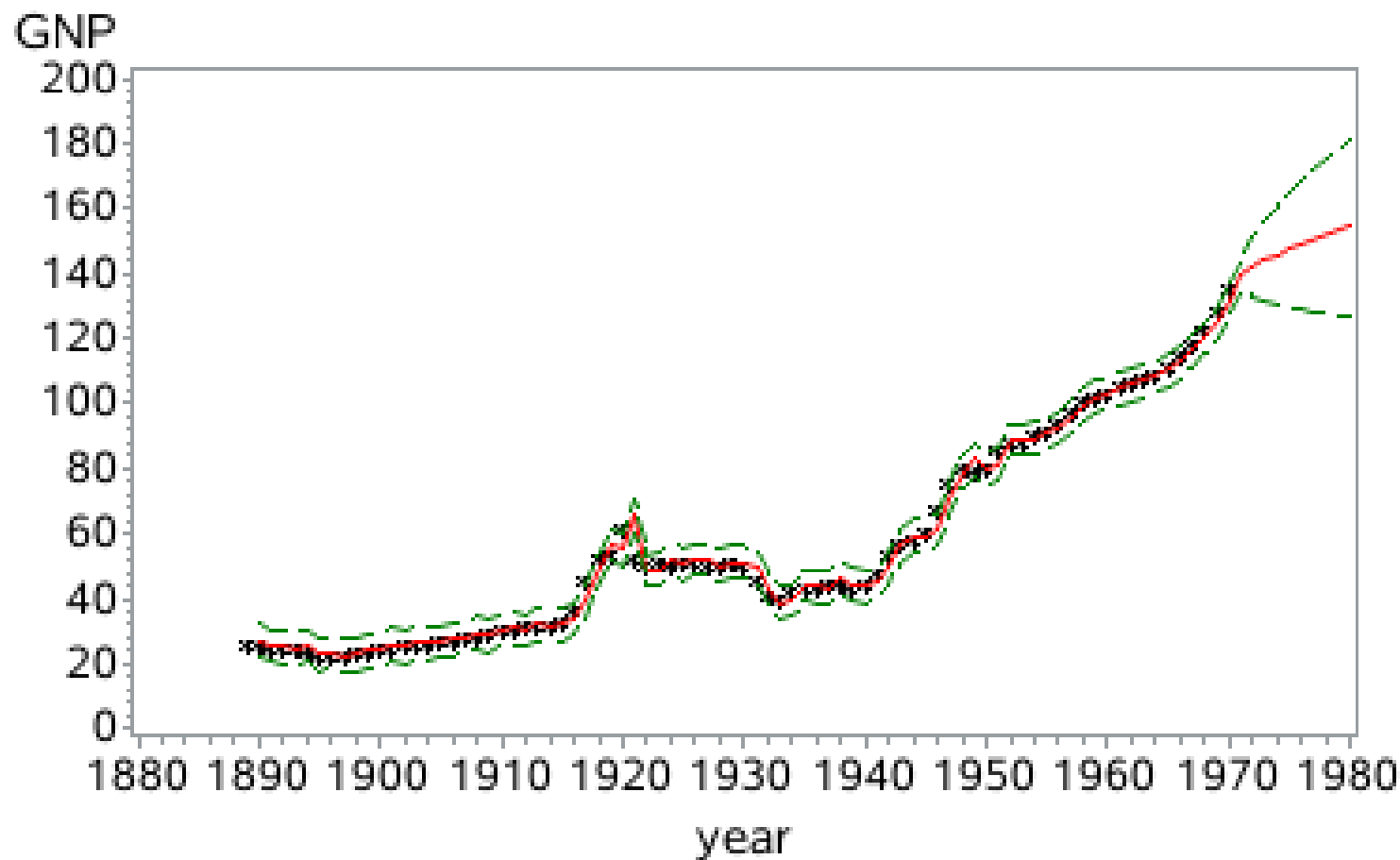
# 例5-6续

- 对1889—1970年美国国民生产总值平减指数序列做为期 1 0 年的预测

年份	预测值	标准差	95%置信下限	95%置信上限
1971	139.362 6	2.521 2	134.421 2	144.304
1972	141.999 1	4.481 3	133.216	150.782 3
1973	143.966 2	6.183 3	131.847	156.085 3
1974	145.618 9	7.660 1	130.605 4	160.632 4
1975	147.124	8.957 8	129.567	164.681 1
1976	148.559 9	10.116 6	128.731 7	168.388 1
1977	149.963 3	11.167 1	128.076 2	171.850 3
1978	151.351 4	12.131 8	127.573 4	175.129 3
1979	152.732 3	13.027 5	127.198 9	178.265 7
1980	154.109 8	13.866 4	126.932 2	181.287 4

## 例5-6拟合与预测效果图

- 对1889—1970年美国国民生产总值平减指数序列做为期10年的预测



# 本章内容

01

**Cramer分解定理**

02

**差分平稳**

03

**ARIMA模型**

04

**疏系数模型**

## 疏系数模型

- ARIMA(p,d,q)模型是指d阶差分后自相关最高阶数为p，移动平均最高阶数为q的模型，通常它包含p+q个独立的未知系数：

$$\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$$

- 如果该模型中有部分自相关系数或部分移动平滑系数为零，即原模型中有部分系数省缺了，那么该模型称为疏系数模型。

## 疏系数模型类型

- 如果只是自相关部分有省缺系数，那么该疏系数模型可以简记为

- $p_1, \dots, p_m$  为非零自相关系数的阶数

$$ARIMA((p_1, \dots, p_m), d, q)$$

- 如果只是移动平滑部分有省缺系数，那么该疏系数模型可以简记为

- $q_1, \dots, q_n$  为非零移动平均系数的阶数

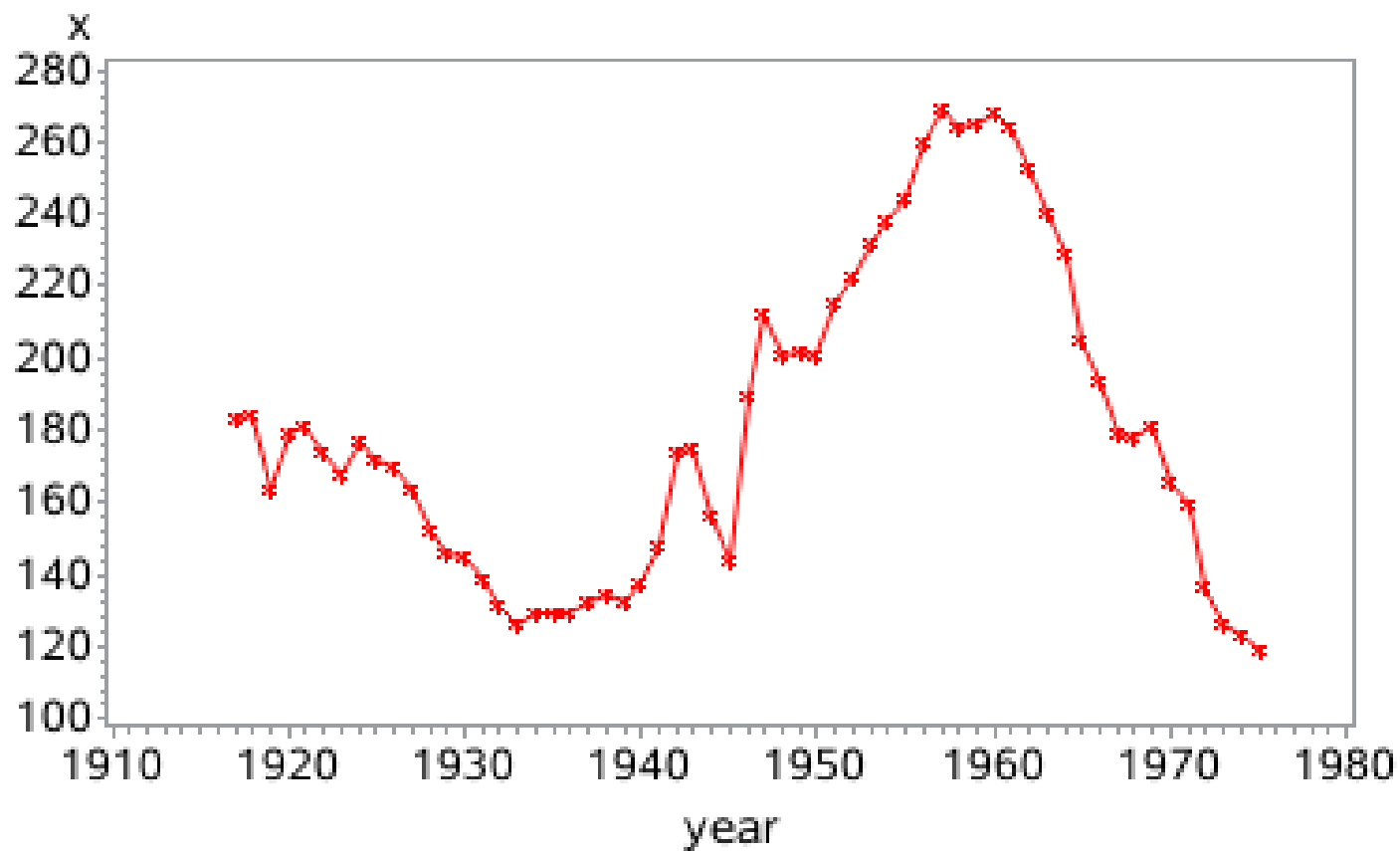
$$ARIMA(p, d, (q_1, \dots, q_n))$$

- 如果自相关和移动平滑部分都有省缺，可以简记为

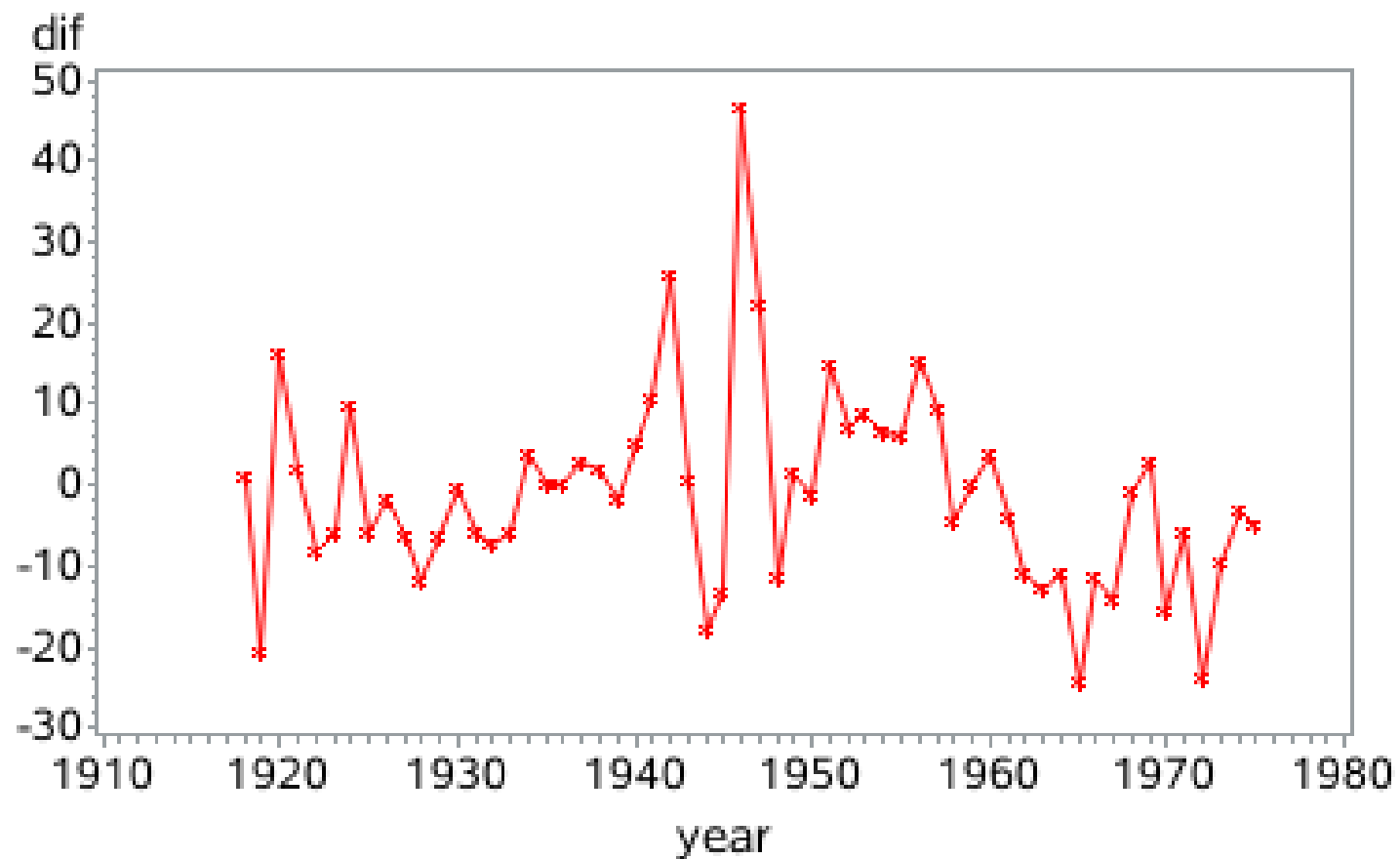
$$ARIMA((p_1, \dots, p_m), d, (q_1, \dots, q_n))$$

## 例5-8

- 对1917年 - 1975年美国23岁妇女每万人生育率序列建模



# 一阶差分

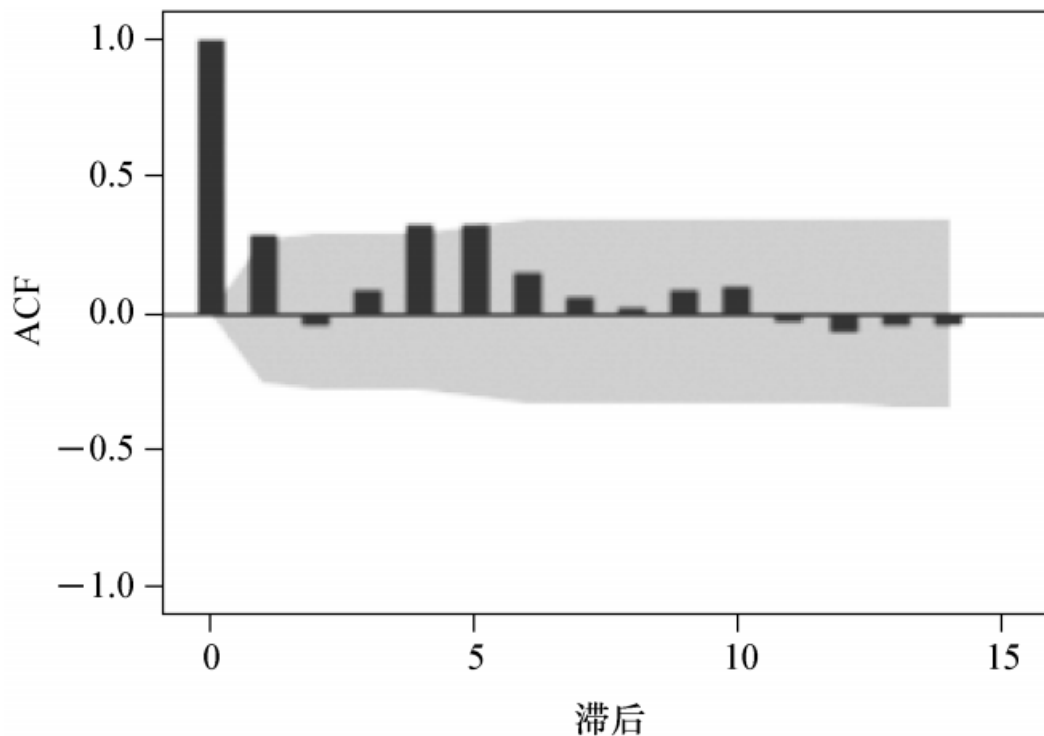


检验结果显示一阶差分后序列为平稳非白噪声序列

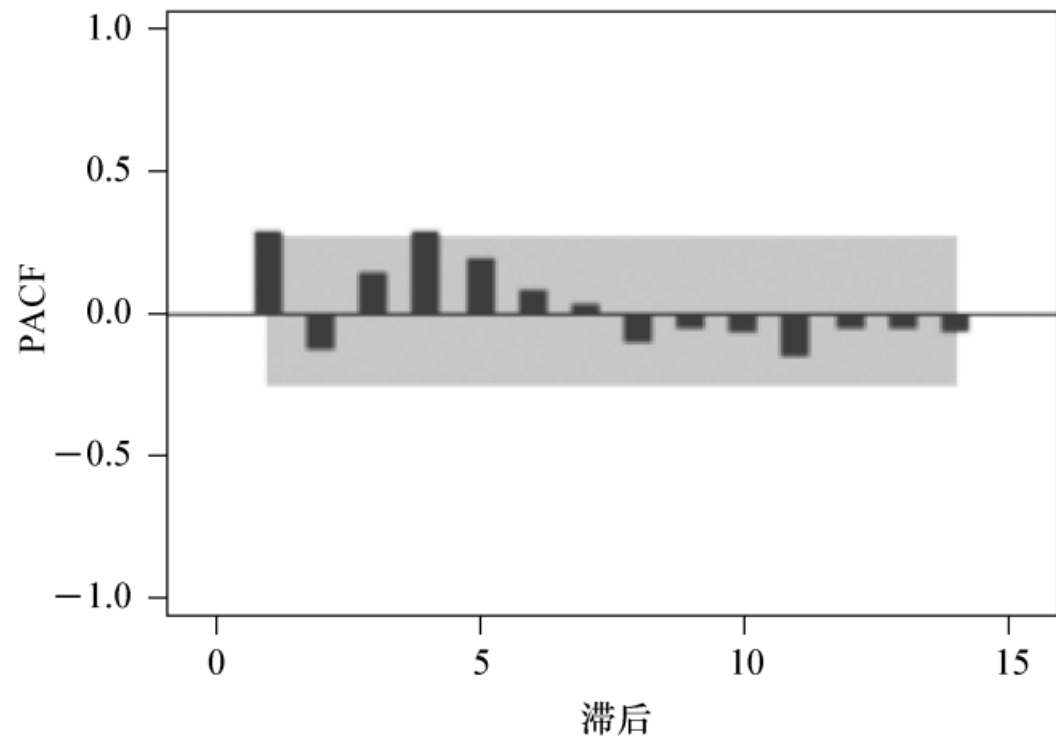


# 模型定阶

- 一阶差分后序列自相关图



- 一阶差分后序列偏自相关图



自相关图显示，延迟 1 阶、4 阶和 5 阶自相关系数大于 2 倍标准差。  
偏自相关图显示，延迟 1 阶和 4 阶的偏自相关系数大于 2 倍标准差。  
根据自相关图和偏自相关图定阶可以有多种尝试。其中一种尝试是认为偏自相关系数 4 阶截尾，可以考虑构建系数 AR(1,4) 模型。

# 模型拟合

- 定阶
  - ARIMA((1,4),1,0)
- 参数估计

$$\nabla x_t = -1.455 + \frac{\epsilon_t}{1 - 0.261B - 0.333B^4}$$

或等价表示为：

$$x_t = -0.591 + 1.261x_{t-1} - 0.261x_{t-2} + 0.333x_{t-4} - 0.333x_{t-5} + \epsilon_t$$

式中，  $\text{Var}(\epsilon_t) = 125.6745$

# 模型检验

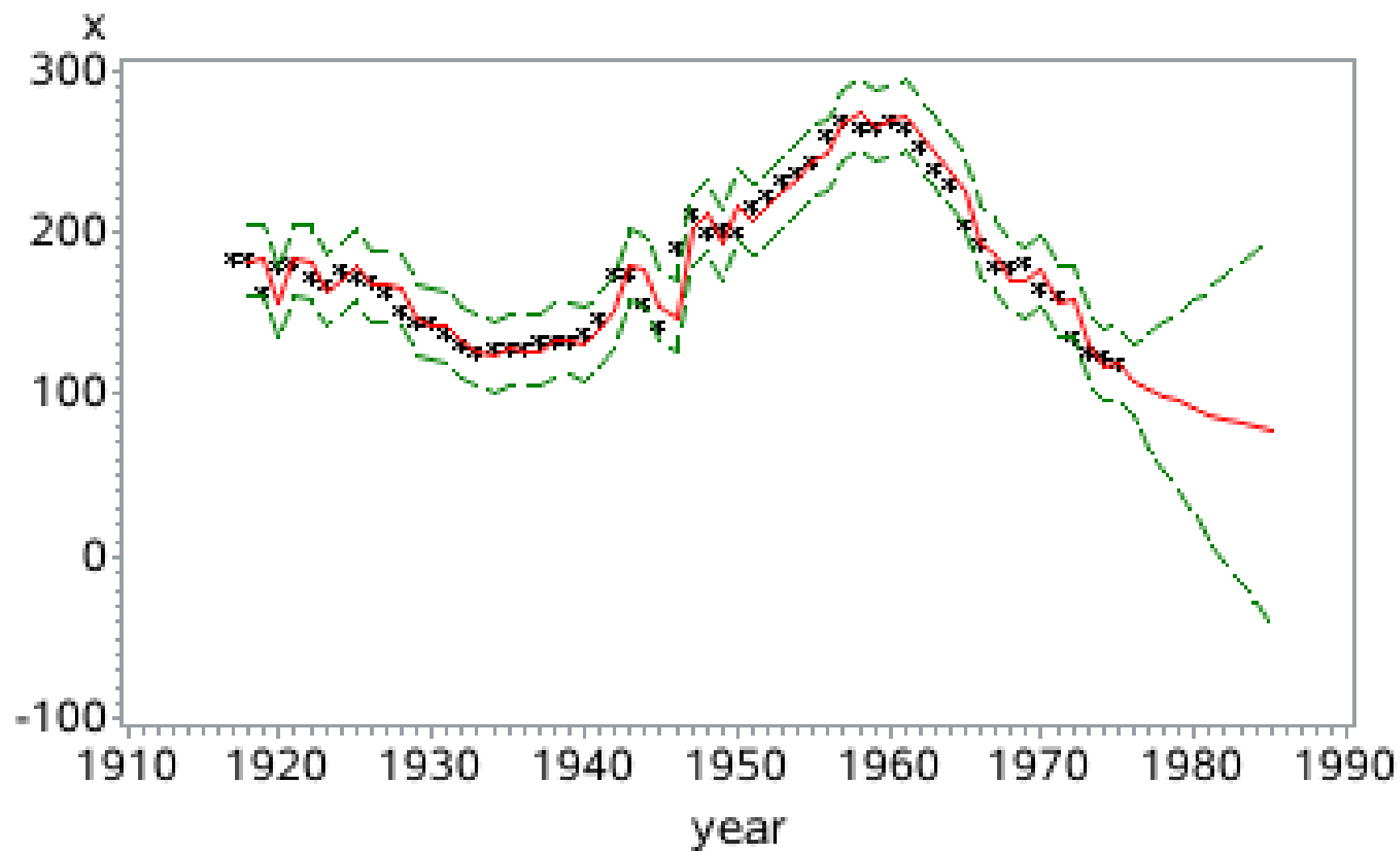
- 模型显著性检验

残差白噪声检验		
延迟阶数	$\chi^2$ 统计量	$P$ 值
6	4.33	0.363 2
12	5.75	0.835 9

- 参数显著性检验

参数显著性检验		
待估参数	$t$ 统计量	$P$ 值
$\phi_1$	2.13	0.037 3
$\phi_4$	2.62	0.011 2

## 拟合预测效果图



**THANKS**

05