条件异方差模型

07

本章内容

方差齐性假定的重要性

- 我们在前面介绍的模型拟合方法（ARIMA模型，因素分解模型）都属于对序列均值的拟合方法

$$\hat{x}_{t+1}=E(x_{t+1})$$

- 但均值的估计值只是一个点估计。对于预测而言，只知道一个点估计没有意义，因为未来真实值恰好等于点估计值的概率近似为 0 。真正有意义的是预测值的置信区间。所以序列预测时，我们更在意预测值的置信区间。之前求的置信区间都基于一个默认的假定——残差序列方差齐性。

$$(\hat{x}_{t+1}-1.96\hat{\sigma}_{t+1}\ ,\ \hat{x}_{t+1}+1.96\hat{\sigma}_{t+1})=(\hat{x}_{t+1}-1.96\hat{\sigma}_{\varepsilon}\ ,\ \hat{x}_{t+1}+1.96\hat{\sigma}_{\varepsilon})$$

- 如果残差序列不满足方差齐性假定，那么置信区间的真实置信度将受到影响。

案例回顾

- 例 5 - 6 中我们对1889-1970年美国GNP平减指数序列进行拟合和预测，得到的拟合模型是：

$$x_t = 0.7292 + 1.4695 x_{t-1} - 0.4695 x_{t-2} + \varepsilon_t \quad , \quad \hat{\sigma}_\varepsilon^2 = 6.3564$$

- 我们很容易预测出1971年美国GNP平减指数的点估计为

$$\hat{x}_{1971} = 139.36$$

- 我们用如下方式计算它的95%置信区间

$$(\hat{x}_{1971} - 1.96\hat{\sigma}_\varepsilon \ , \ \hat{x}_{1971} + 1.96\hat{\sigma}_\varepsilon) = (134.42, 144.30)$$

- 这个95%的置信区间一定有95%的置信水平吗？不一定，它的准确性取决于它是否满足如下假定

$$\hat{\sigma}_{1971}^2 \overset{?}{=} \hat{\sigma}_\varepsilon^2$$

白噪声检验没有包含方差齐性检验

- 在进行ARIMA模型拟合时，我们最后是通过检验残差序列是否通过白噪声检验，来判断模型的显著性。
- 白噪声序列应该满足三个条件：
  - 零均值 $E(\varepsilon_t)=0$
  - 纯随机 $Cov(\varepsilon_t,\varepsilon_{t-i})=0 \quad ,\forall i\geq 1$
  - 方差齐性 $Var(\varepsilon_t)=\sigma^2$
- 实际上，我们构造的白噪声检验统计量（Q统计量）只检验了白噪声序列的第二个条件。第一个条件在建模时，常数项拟合可以满足（可以不检验）。
- 也就是说，白噪声检验只检验了其中两条假定条件。第三条假定：方差齐性是没有检验的，我们是默认它满足的。但实际上，这个条件并不总是满足。

异方差属性

- 如果残差序列方差齐性的假定不成立，即随着时间的变化，残差序列的方差不是常数，我们称这种属性为方差非齐，或简称为异方差现象。序列的异方差属性可以表达为方差为时间 t 的函数：

$$Var(\varepsilon_t) = \sigma_t^2$$

- 如果忽视异方差的存在，就会使得残差的方差估计不准确。相应地，置信水平为 1 -α的置信区间，实际的置信水平就不是1-α了。这会使得估计和预测的精度都受到影响。更糟糕的是，会受到多大的影响，我们都无法测量。所以为了提高拟合模型的估计和预测精度，我们需要对残差序列进行方差齐性检验，并对异方差序列进行深入的分析。

异方差的直观诊断

- 有些序列具有明显的异方差属性，通过残差图或者残差平方图就可以直观看出来

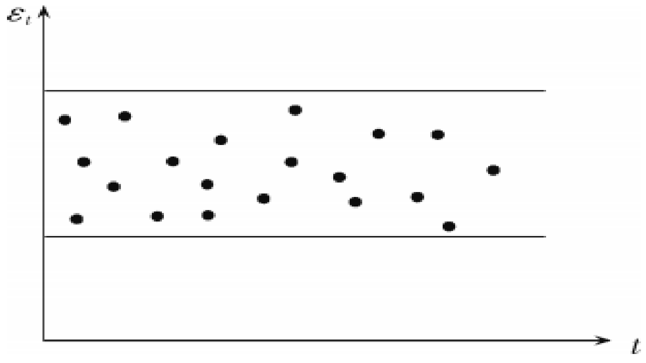


图 7－1　方差齐性残差图

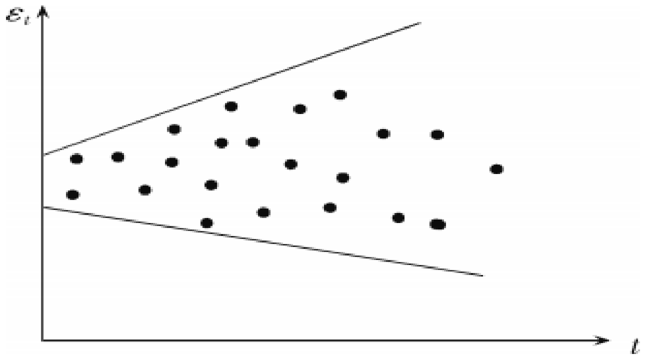


图 7－2　递增型异方差

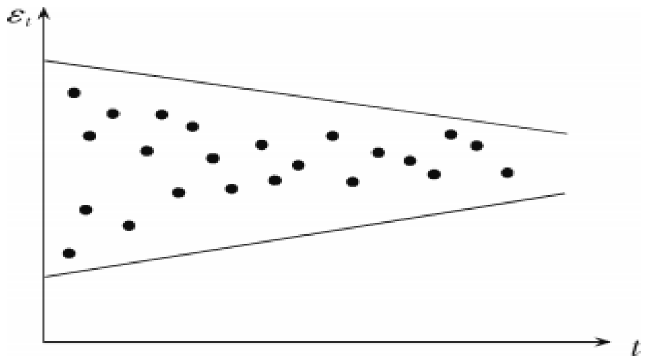


图 7－3　递减型异方差

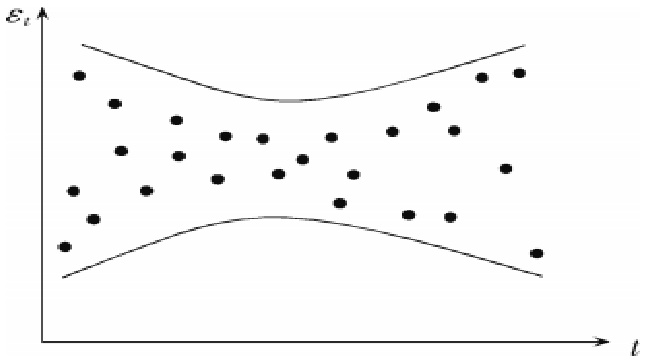


图 7－4　综合型异方差

残差平方图

- 由于残差序列的方差实际上就是它平方的期望，即

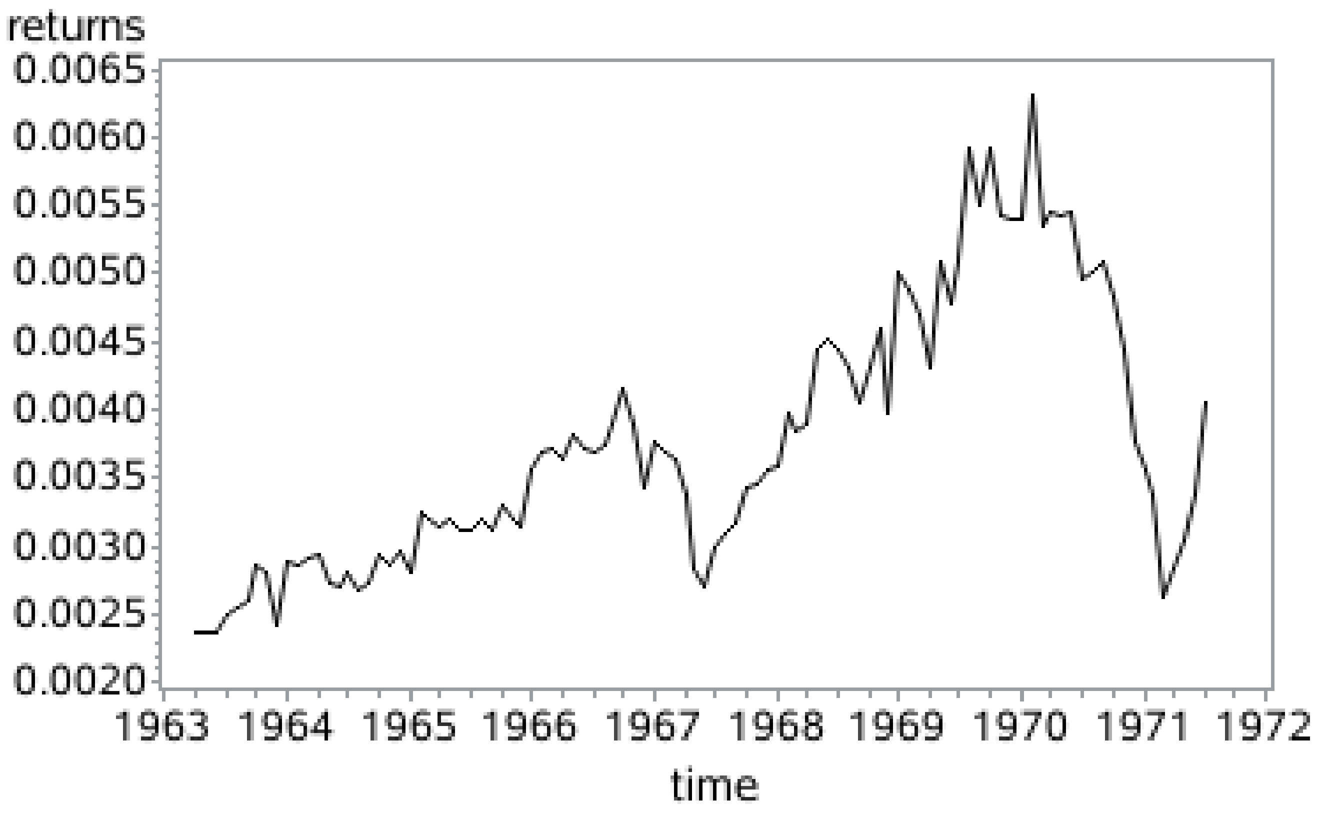
$$\mathrm{Var}(\varepsilon_t)=E(\varepsilon_t^2)$$

- 所以残差序列是否方差齐性，主要是考察残差平方序列的性质。我们可以借助残差平方图对残差序列的方差齐性进行直观诊断
- 假设方差齐性满足，则残差平方序列应该在某个常数值附近随机波动，它不应该具有任何明显的规律性，否则就呈现出异方差属性

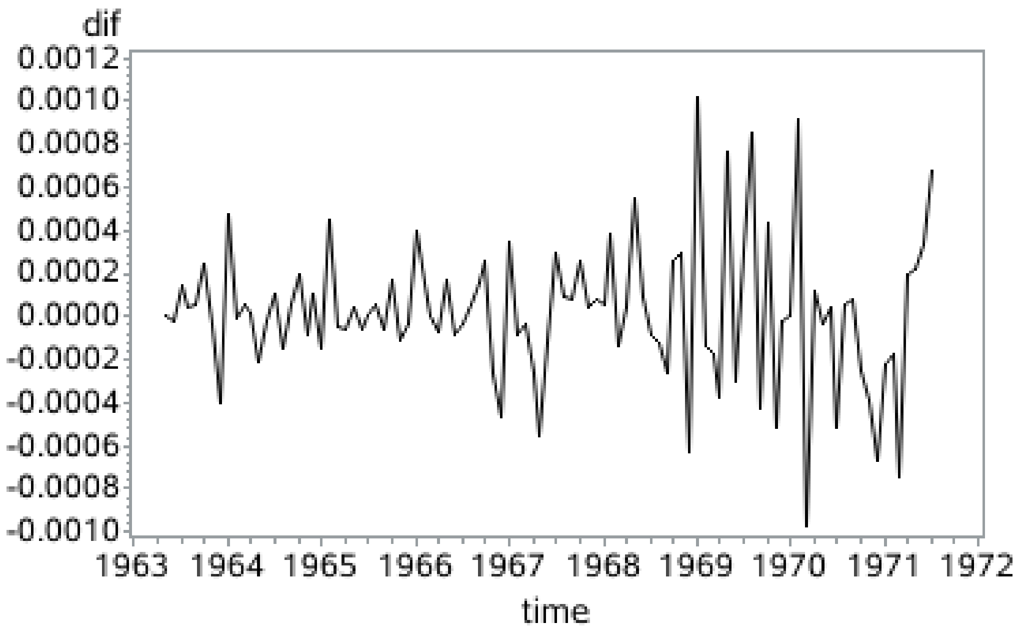
$$E(\varepsilon_t^2)=\sigma_\varepsilon^2$$

例7-1

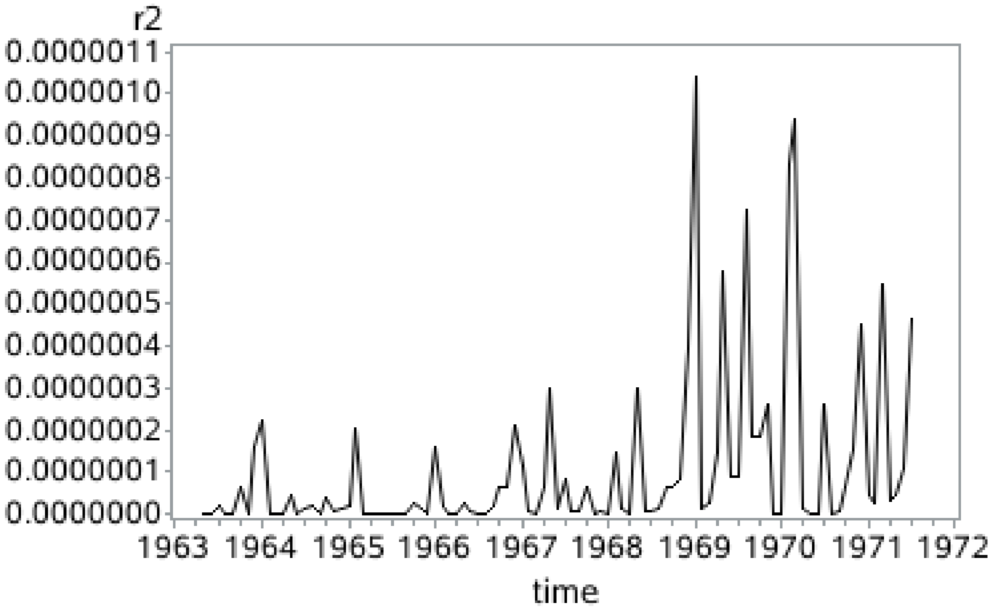
- 直观考察1963年 4 月至1971年 7 月美国短期国库券的月度收益率序列的方差齐性



方差非齐特征



1阶差分后残差序列时序图



1阶差分后残差平方序列时序图

异方差处理方法

- 目前主要有两种异方差的处理方法：
  - 方法一：假如已知异方差函数的具体形式，找到它的转换函数，进行方差齐性变换

$$\sigma_t^2 = h(\mu_t)$$

$$g'(x) = \left(h(x)\right)^{-1/2}$$

$$Var[g(x_t)] = \sigma^2$$

方法二：假如异方差显示出集群效应，拟合条件异方差模型

$$\sigma_t^2 = f(\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \varepsilon_{t-3}^2, \cdots)$$

本章内容

方差齐性变换思想

- 使用场合：假设序列显示出显著的异方差性，且方差与均值之间具有某种明确的函数关系

$$\sigma_t^2 = h(\mu_t)$$

其中，$h(\cdot)$ 是某个已知函数。

- 在这种场合下，我们的处理思路是尝试寻找一个转换函数 g (·) ，使得经转换后的变量满足方差齐性

$$Var[g(x_t)] = \sigma^2$$

常用的异方差变换

- 在实践中，许多金融时间序列都呈现出异方差的性质，而且通常序列的标准差与其水平均值之间具有某种正比关系，即序列的水平均值小时，序列的波动范围小，序列的水平均值大时，序列的波动范围大。对于这种异方差性质，最简单的假定为

$$\sigma_t = \mu_t$$

- 等价于

$$\sigma_t^2 = h(\mu_t) = \mu_t^2$$

- 寻找转换函数g()，使得转换后函数满足方差齐性

$$Var\left[g(x_t)\right] = \sigma^2$$

转换函数的确定

- 将 $g(x_t)$ 在 μ_t 附近做1阶泰勒展开

$$g(x_t) \approx g(\mu_t) + (x_t - \mu_t) g'(\mu_t)$$
$$\Rightarrow Var\left[g(x_t)\right] \approx \left[g'(\mu_t)\right]^2 Var\left(x_t\right)$$
$$\Rightarrow \left[g'(\mu_t)\right]^2 h\left(\mu_t\right) = \sigma^2$$

- 这时方差齐性的变换函数为

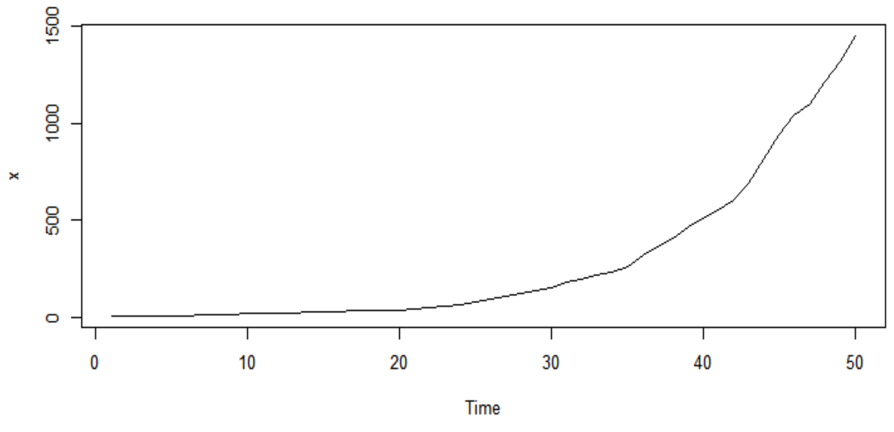
$$g'(\mu_t) = \frac{1}{\sqrt{h(\mu_t)}} = \frac{1}{\mu_t} \quad \Rightarrow \quad g(\mu_t) = \ln(\mu_t)$$

- 这意味着对于标准差与水平均值成线性正比关系时的异方差序列，对数变换可以实现方差齐性变换。

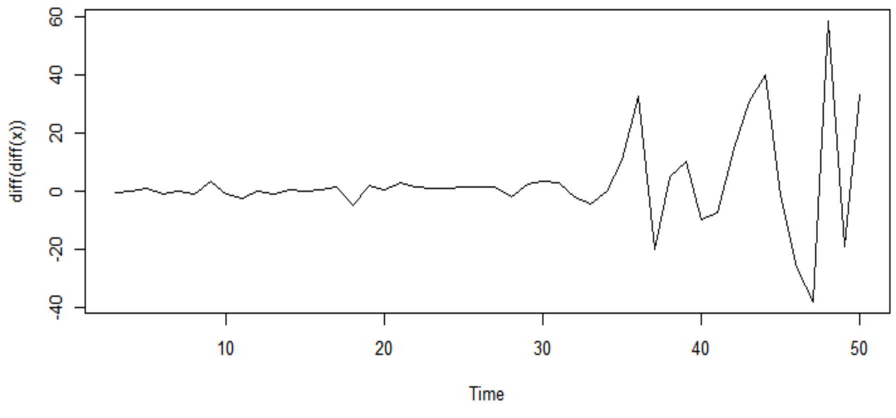
例5-2续

- 对1950-1999年北京市民用车辆拥有量序列的异方差性进行考察，并进行方差齐性变换

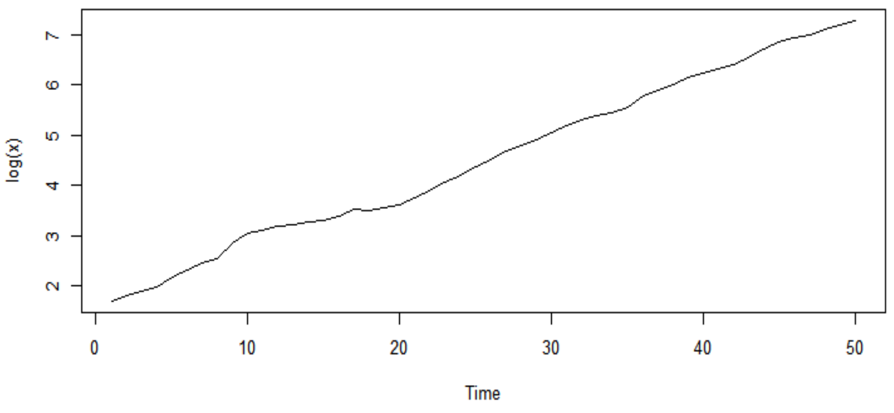
原序列时序图



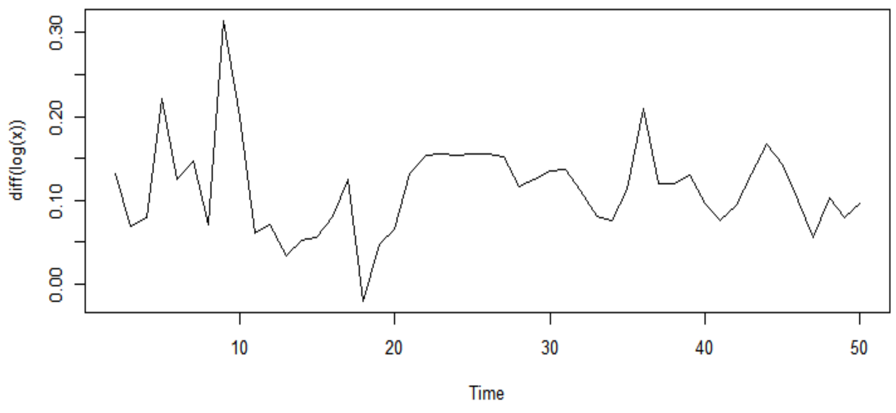
原序列2阶差分后时序图



对数序列时序图



对数序列1阶差分后时序图



异方差变换的普适性和局限性

- 普适性
  - 由于很多经济和金融变量都具有方差随着均值递增而递增的特点，所以在实务领域，经济学家和金融研究人员都会在建模之前先对序列进行对数变换，希望能消除方差非齐。
- 局限性
  - 残差序列的方差与原序列均值之间的关系非有各种可能，不一定就是线性递增关系。所以并不是所有序列都能使用对数变换进行异方差信息提取。
  - 没有办法确定序列的方差函数与均值函数之间的函数关系。所以异方差变换通常只能是明显的线性递增型异方差的预处理手段，并不是一种精准科学的建模方法。
  - 异方差变换不是提取波动性信息的主流方法。

本章内容

波动性分析产生的背景

- 1982年，Engle根据1958年2季度至1977年2季度的数据，研究英国因工资上涨导致通货膨胀问题时，对物价指数序列构建了一个AR(4)自回归模型

$$
\begin{aligned}
P_t &= 0.0257 + 0.334P_{t-1} + 0.408P_{t-4} - 0.404P_{t-5} \\
&\quad - 0.0559(P_{t-1} - W_{t-1}) + \varepsilon_t
\end{aligned}
$$

其中：P是物价指数，W是工资水平， $Var(\varepsilon_t) = 23\times 10^{-6}$

- 在方差齐性的假定下，向前做1期预测，很容易预测出1977年3季度物价指数的95%的波动范围为

$$
(\hat{P}_{t+1} - 1.96\sqrt{23\times 10^{-6}}, \hat{P}_{t+1} + 1.96\sqrt{23\times 10^{-6}})
$$

波动性分析产生的背景

- 但是Engle以经济学家的经验，认为这个预测的置信区间偏小，与实际情况严重不符。因为从1974年开始物价指数的平均波动等于

$$230\times10^{-6}$$

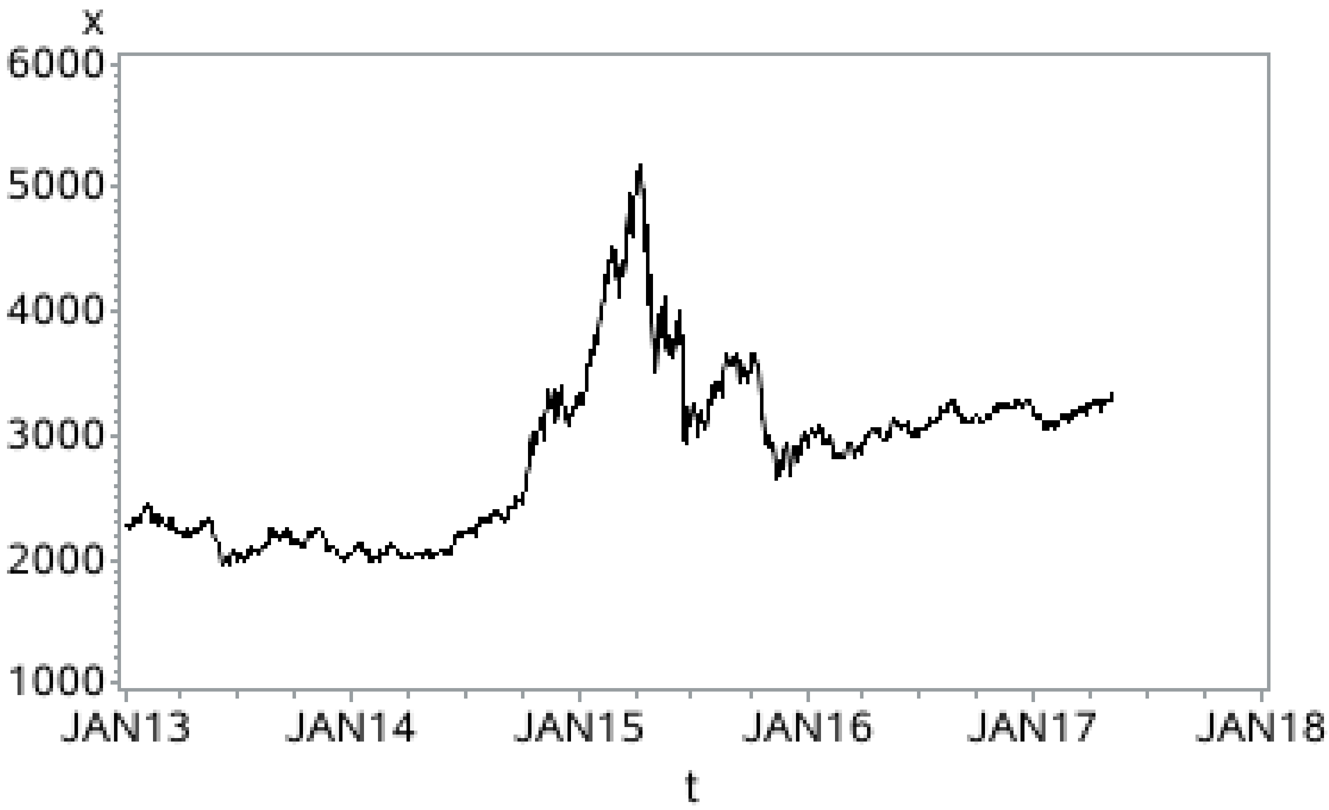
- 也就是说物价指数最近4年的方差是过去20年方差的10倍。近期没有任何宏观政策公布以压制物价波动，物价指数序列也没有任何波动减缓的趋势。按照正常的经济运行惯性，下一季度（1977年3季度）的物价指数波动不可能突然下降10倍。
- Engle认为，按照方差齐性假定求出的1977年3季度物价指数波动的95%置信区间，一定没有达到95%的置信程度。
- 那真实的波动信息蕴含在哪呢？更准确的置信区间该怎么求呢？基于Engle的研究，如果残差序列具有集群效应，我们可以使用ARCH模型得到更准确的置信区间。

集群效应

- 所谓集群效应（volatility cluster）是指在消除确定性非平稳因素的影响后，残差序列在大部分时段小幅波动，但是会在某些时段出现持续大幅波动。于是序列的波动就呈现出一段持续时间的小幅波动和一段持续时间的大幅波动交替出现的特征。
- 集群效应是很多经济和金融序列都具有的波动特征。1963年，Benoit Mandelbrot就指出：在金融市场中数据通常比正态分布存在更多异常值，且具有集群效应。
- 集群效应的产生原因，通常认为是经济市场和金融市场的波动易受谣言、政局变动、政府货币与财政政策变化等诸多因素的影响
  - 一旦某个影响因素出现，市场会大幅波动，以消化这个影响，这就出现密集的大幅波动。
  - 波动到位实现新的稳定之后，在下一个影响因素到来之前，序列会维持一段时间的小幅波动

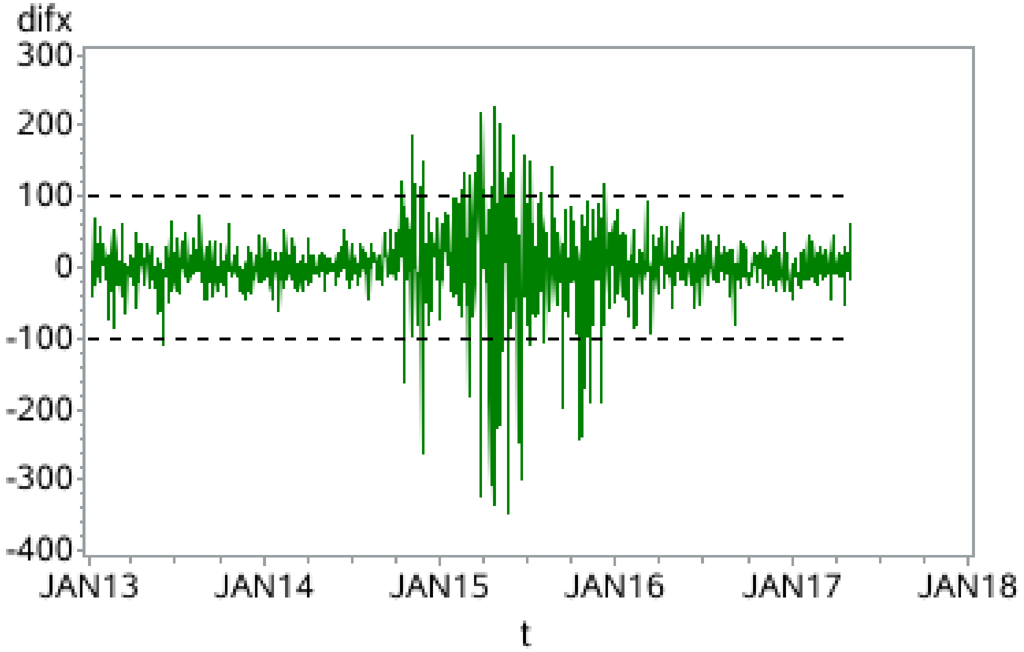
例7-2

- 考察2013年 1 月 4 日至2017年8月25日上证指数每日收盘价序列的集群效应特征。



一阶差分后序列具有显著的集群效应特征

- 右图显示上证指数1阶差分后残差序列在2015年之前序列持续小幅波动，2015-2016年序列持续大幅波动，2016年之后又持续小幅波动，这就是显著的集群效应特征。
- 集群效应的存在意味着在序列持续大幅波动的时期，基于方差齐性假定得到的95%置信区间会显著小于序列的真实波动范围。所谓95%的置信区间，在2015-2016年根本达不到95%的置信水平。



上证指数1阶差分后序列时序图
图中虚线为方差齐性假定下95%置信区间

ARCH模型的构造思想

- ARCH模型的构造思想
  - Engle认为集群效应的特征是一段时间小幅波动，再一段时间大幅波动，这意味着序列的波动存在相关性。因为如果序列的波动不存在相关性的话，就不会产生小幅波动和大幅波动集中交替出现，而是呈现出大幅波动和小幅波动完全无规律。
  - 基于这个思想，Engle构造了自回归条件异方差模型（autoregressive conditional heteroskedastic model ，简记为ARCH模型）
- 最早的ARCH模型
  - Engle重新拟合1958年2季度至1997年2季度英国物价指数序列

$$
\begin{cases}
P_t = 0.0321 + 0.021P_{t-1} + 0.27P_{t-4} - 0.334P_{t-5} - 0.0697(P_{t-1} - W_{t-1}) + \varepsilon_t \\
Var(\varepsilon_t) = h_t \quad \text{------------ 方差非齐} \\
h_t = 19\times 10^{-6} + 0.846(0.4\varepsilon_{t-1}^2 + 0.3\varepsilon_{t-2}^2 + 0.2\varepsilon_{t-3}^2 + 0.1\varepsilon_{t-4}^2) \quad \text{------------ ARCH(4)}
\end{cases}
$$

ARCH模型的结构

- 零均值残差序列具有异方差属性

$$E(\varepsilon_t)=0,\quad Var(\varepsilon_t \mid \varepsilon_{t-1},\varepsilon_{t-2},\cdots)=h_t$$

- 条件方差具有短期线性相关性，于是可以将历史波动信息作为条件，采用线性回归的方式估计序列的当期波动

$$h_t=\lambda_0+\sum_{i=1}^{q}\lambda_i\varepsilon_{t-i}^2$$

具有上面结构的模型称为q阶自回归条件异方差模型，简记为ARCH(q)。

ARCH模型的参数约束条件

（1）因为方差必须是非负的，即

$$Var(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \cdots) \geq 0 \Rightarrow \lambda_0 + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2 \geq 0$$

要保证对任意 ε_{t-i} 上式都要成立，只能要求每个参数都非负 $\lambda_0, \lambda_1, \cdots, \lambda_q \geq 0$

（2）要保证无条件方差 $Var(\varepsilon_t) = \sigma^2$ 存在且方差不能为负

$$E\left[Var(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \cdots)\right] = E\left(\lambda_0 + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2\right) \Rightarrow \sigma^2 = \frac{\lambda_0}{1 - \sum_{i=1}^{q} \lambda_i}$$

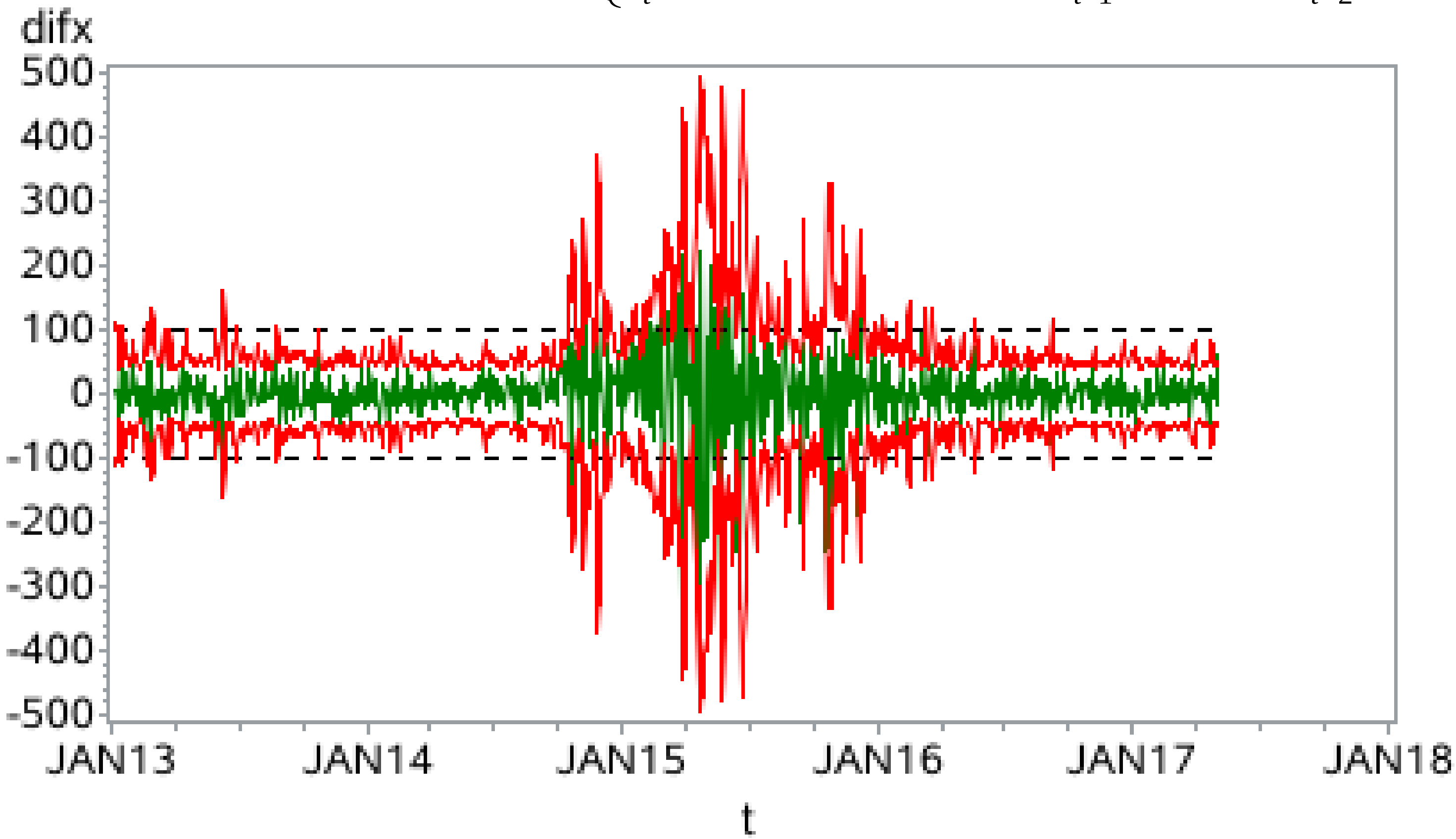
要保证无条件方差 σ^2 存在且方差不能为负，就要求上式分母大于零

$$1 - \sum_{i=1}^{q} \lambda_i > 0 \Rightarrow \sum_{i=1}^{q} \lambda_i < 1$$

例7-2续

- 对2013年1月4日至2017年8月25日上证指数每日收盘价序列拟合 ARCH 模型，并考察条件异方差和无条件方差的差别。

$$\begin{cases} \varepsilon_t = \sqrt{h_t} e_t \quad, e_t \sim N(0,1) \\ h_t = 383.5301 + 0.0462\varepsilon_{t-1}^2 + 0.4675\varepsilon_{t-2}^2 + 0.2681\varepsilon_{t-3}^2 + 0.2369\varepsilon_{t-4}^2 \end{cases}$$



本章内容

GARCH模型的产生背景

- Engle的学生Tim Bollerslev（1986）对美国1948年2季度-1983年4季度GNP平减指数（GNP Deflator)的对数序列建模，他拟合了如下两个模型
- 模型一：方差齐性模型

$$\pi_t = 0.24 + 0.552\pi_{t-1} 0.177\pi_{t-2} + 0.232\pi_{t-3} - 0.209\pi_{t-4} + a_t$$
$$Var(a_t) = 0.282$$

- 模型二：条件异方差模型

$$\pi_t = 0.138 + 0.423\pi_{t-1} 0.222\pi_{t-2} + 0.377\pi_{t-3} - 0.175\pi_{t-4} + a_t$$
$$a_t = \sigma_t \varepsilon_t$$
$$\sigma_t^2 = 0.058 + 0.802\sum_{i=1}^{8}[(9-i)/36]a_{t-i}^2$$

- 美国GNP平减指数误差平方序列具有长期相关，上述ARCH(8)模型都不能将波动信息提取充分

ARCH模型的短期相关属性

- ARCH 模型的实质是使用残差平方序列的 q 阶移动平均拟合当期异方差函数值

$$h_t = \lambda_0 + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2$$

- 由于移动平均模型具有自相关系数q阶截尾性，所以ARCH模型实际上只适用于异方差函数短期自相关过程。
- 但是在实践中，有些残差序列的异方差函数是具有长期自相关性的，这时如果使用ARCH模型拟合异方差函数，将会产生很高的移动平均阶数。这会增加参数估计的难度并最终影响ARCH模型的拟合精度。
- Bollerslev基于这种考虑，提出了广义自回归条件异方差模型（generalized autoregressive conditional heteroskedastic model），简记为GARCH（1，1）模型

最早的GARCH模型

- Bollerslev（1986）对ARCH模型进行改进，引入异方差的历史信息，拟合了如下形式的条件异方差模型

$$\begin{cases} \pi_t = 0.141 + 0.433\pi_{t-1} + 0.229\pi_{t-2} + 0.349\pi_{t-3} - 0.162\pi_{t-4} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t, e_t \sim N(0,1) \\ h_t = 0.007 + 0.135 h_{t-1} + 0.829\varepsilon_{t-1}^2 \end{cases}$$

- 为了更清楚地理解GARCH模型的实质，把异方差序列$\{h_t\}$视作为响应序列，把残差平方序列$\{\varepsilon_t^2\}$视作为随机扰动项序列，那么ARCH模型实际上就是$\{h_t\}$关于$\{\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \cdots, \varepsilon_{t-q}^2\}$的q阶移动平均的模型。而GARCH模型实际上就是$\{h_t\}$关于$\{h_{t-1}, \cdots, h_{t-p}\}$的p阶自相关，关于$\{\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \cdots, \varepsilon_{t-q}^2\}$的q阶移动平均的模型。
- 显然ARCH模型是GARCH模型的一个特例。

具有条件异方差属性的单变量序列完整的分析结构

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \cdots) + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t \quad , \quad e_t \overset{i.i.d}{\sim} N(0,1) \\ h_t = \lambda_0 + \sum_{i=1}^{p} \eta_i h_{t-i} + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2 \end{cases}$$

均值模型

异方差属性，分布假定

条件异方差模型

一个完整的条件异方差模型，就是由这三部分：均值模型，条件异方差模型和分布假定构成的。

拟合GARCH模型的步骤

- 拟合GARCH模型的步骤如下：

  第一步：构建水平模型，提取序列均值中蕴涵的相关信息。

  第二步：检验残差序列是否具有条件异方差特征。

  第三步：对具有条件异方差特征的序列拟合GARCH模型。

  第四步：检验拟合模型的优劣，优化模型。

  第五步：使用拟合模型进行预测。

PP检验

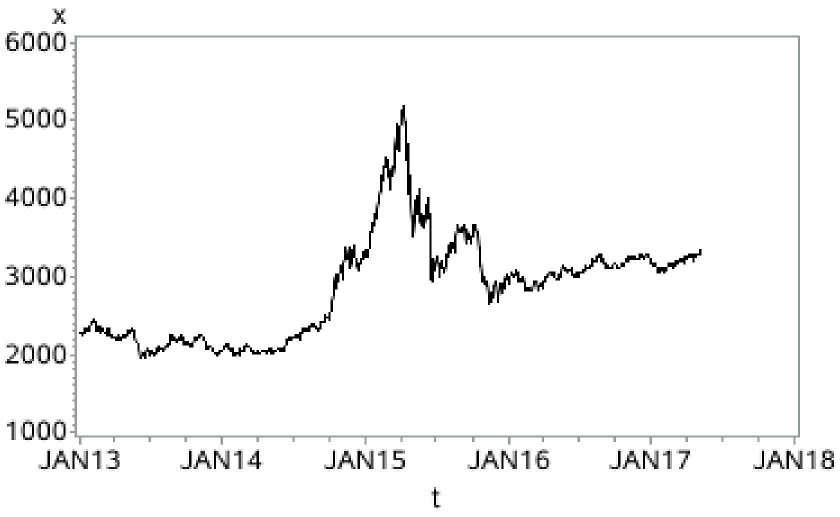
- 建立条件异方差模型，首先需要提取序列的均值 （确定性）信息，而最常使用的均值模型是 ARIMA模型。要建立 ARIMA模型，必须首先对序列的平稳性进行判断。 而ADF检验是在方差齐性假定下构造的平稳性检验统计量。它对异方差序列的平稳性检验可能会有偏差。
- Phillips和 Perron在1988年对 ADF检验进行了非参数修正，提出了Phillips-Perron检验统计量，简称为PP检验。该检验统计量适用于异方差场合的平稳性检验。
- PP统计量的构造

$$Z(\tau)=\tau(\hat{\sigma}^2/\hat{\sigma}_{Sl}^2)-(1/2)(\hat{\sigma}_{Sl}^2-\hat{\sigma}^2)T\sqrt{\hat{\sigma}_{Sl}^2\sum_{t=2}^{T}(x_{t-1}-\bar{x}_{T-1})^2}$$

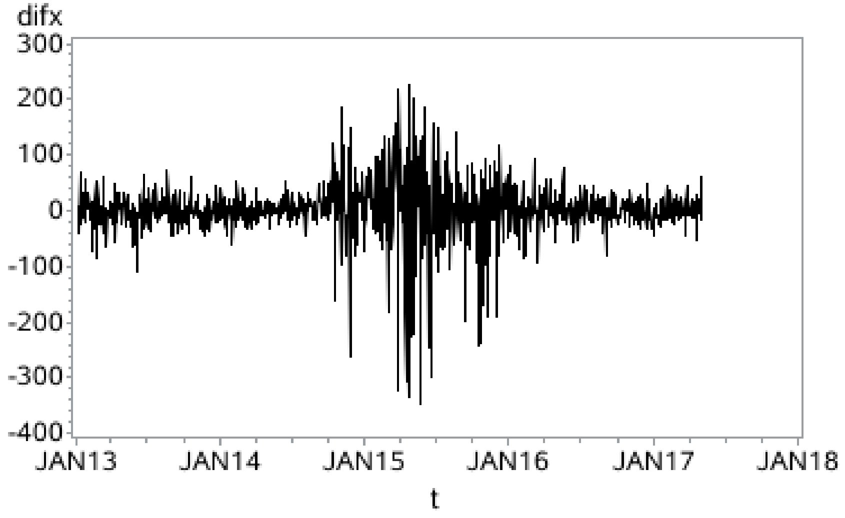
- 修正后的 $Z(\tau)$ 统计量和τ统计量具有相同的极限分布。这就意味着对于异方差序列只需要在原来τ统计量的基础上进行一定的修正，构造出 $Z(\tau)$ 统计量。$Z(\tau)$ 统计量不仅考虑到自相关误差所产生的影响，还可以继续使用τ统计量的临界值表进行检验，而不需要拟合新的临界值表。

例7-2续

- 对2013年 1 月 4 日至2017年 8 月25日上证指数每日收盘价序列提取均值信息.



上证指数序列时序图



上证指数1阶差分后序列时序图

该序列时序图呈现出典型的先递增后递减的非平稳趋势特征。使用 1 阶差分运算提取该序列的确定性趋势信息。差分后序列呈现出典型的集群效应。考虑方差非齐的影响，对差分后序列平稳性的检验，最好使用PP检验。

差分后序列PP检验

- 上证指数1阶差分后序列PP检验

类型	延迟阶数	τ统计量的值	$Pr<\tau$
类型一	0	−25.37	<0.000 1
	1	−25.37	<0.000 1
	2	−25.36	<0.000 1
类型二	0	−25.61	<0.000 1
	1	−25.61	<0.000 1
	2	−25.59	<0.000 1
类型三	0	−25.6	<0.000 1
	1	−25.61	<0.000 1
	2	−25.59	<0.000 1

- PP检验结果：1阶差分后序列平稳。

异方差序列纯随机性检验

- 传统的纯随机性检验都是借助LB检验统计量进行的，而LB检验统计量是在序列满足方差齐性的假定下构造的。当序列存在异方差属性时，LB统计量不再近似服从卡方分布。也就是说，在条件异方差存在的场合，白噪声检验结果可能不再准确。
- 通常出现的问题就是残差序列之间的相关系数已经很小，近似白噪声序列，但是白噪声检验结果却显示LB检验统计量的 P 值很小。这时不能单纯依据LB检验统计量的结果做出判断。
- 在异方差可能存在的场合，LB检验结果只能作为参考信息之一，同时还要参考自相关系数的大小，如果自相关系数都很小 （比如都小于０.２），可以认为序列近似白噪声序列。

例7-2续

- 上证指数每日收盘价1阶差分序列，进行纯随机性检验。
- 白噪声检验结果

白噪声检验									
延迟	LB	自由度	P 值	自相关系数					
6	65.04	6	<.0001	0.089	-0.095	0.013	0.171	0.014	-0.104
12	104.02	12	<.0001	0.03	0.125	0.042	-0.106	-0.06	0.033
18	153.58	18	<.0001	0.088	-0.154	0.043	0.083	0.026	-0.05
24	199.16	24	<.0001	-0.021	0.107	0.097	-0.075	-0.112	-0.007

- 白噪声检验结果：

例7-2续：差分后序列纯随机性检验

- 上证指数1阶差分后序列纯随机性检验

白噪声检验									
延迟	LB	自由度	P 值	自相关系数					
6	65.04	6	<.0001	0.089	-0.095	0.013	0.171	0.014	-0.104
12	104.02	12	<.0001	0.03	0.125	0.042	-0.106	-0.06	0.033
18	153.58	18	<.0001	0.088	-0.154	0.043	0.083	0.026	-0.05
24	199.16	24	<.0001	-0.021	0.107	0.097	-0.075	-0.112	-0.007

- 白噪声检验结果：LB检验的P值都极小，方差齐性假定下的判断是该序列为非白噪声序列。但本例可能方差非齐，且延迟各阶的自相关系数都很小，最大的 $\rho_4 = 0.171$。所以综合考虑，可以认为差分后序列近似为白噪声序列。
- 所以上证指数的均值模型为ARIMA(0,1,0)模型：$x_t = x_{t+1} + \varepsilon_t$

条件异方差检验

- 提取了均值信息后，需要对残差序列进行条件异方差检验，判断是否需要进一步拟合条件异方差模型
- 条件异方差检验，也称为 ARCH 检验。ARCH 检验是一种特殊的异方差检验，它不仅要求序列具有异方差性，而且要求这种异方差性是由某种自相关关系造成的，这种自相关关系可以用残差序列的自回归模型进行拟合。
- 常用的两种 ARCH 检验统计量
  - 1983年 Mcleod和Li提出了Portmanteau Q统计量
  - 1982年Engle提出的拉格朗日乘子检验，简记为LM 统计量

Portmanteau Q检验

- 构造思想
  - 如果残差序列方差非齐，且具有集群效应，那么残差平方序列通常具有自相关性。所以方差非齐检验可以转化为残差平方序列的自相关性检验。
- 假设条件：

H_0：残差平方序列纯随机 $\leftrightarrow$ H_1：残差平方序列自相关
（方差齐性） （方差非齐）

用 ρ_k 表示残差平方序列 $\{\varepsilon_t^2\}$ 的延迟 k 阶自相关系数，则该假设条件可以等价表达为：

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_q = 0 \leftrightarrow H_1: \rho_1, \rho_2, \cdots, \rho_q \text{ 不全为零}$$

- 检验统计量

$$Q(q) = n(n+2)\sum_{i=1}^{q}\frac{\rho_i^2}{n-i} \sim \chi^2(q-1)$$

LM检验

- 构造思想
  - 假如残差序列方差非齐，且具有集群效应，残差平方序列通常具有自相关性。那么我们就可以尝试使用自回归模型拟合残差平方序列，于是方差齐性检验就可以转化为这个方程是否显著成立的检验

$$\varepsilon_t^2 = \lambda_0 + \sum_{i=1}^{q} \lambda_i \varepsilon_{t-i}^2 + v_t$$

- 假设条件

H_0：残差平方序列纯随机 $\leftrightarrow$ H_1：残差平方序列具有自相关性

假设条件等价为：

$H_0: \lambda_1 = \lambda_2 = \cdots = \lambda_q = 0 \leftrightarrow H_1: \lambda_1, \lambda_2, \cdots, \lambda_q$ 不全为零

- 检验统计量

$$LM(q) = \frac{(SST - SSE)/q}{SSE/t - 2q - 1} \sim \chi^2(q-1)$$

例7-2续

- 对2013年1月4日至2017年8月25日上证指数每日收盘价序列进行ARCH检验。

阶数	Q	$Pr>Q$	LM	$Pr>$LM
1	78.283 1	<0.000 1	78.110 9	<0.000 1
2	200.425 8	<0.000 1	159.622 1	<0.000 1
3	313.732 1	<0.000 1	202.526 8	<0.000 1
4	430.930 0	<0.000 1	231.986 0	<0.000 1
5	506.692 1	<0.000 1	236.765 0	<0.000 1
6	557.562 0	<0.000 1	236.770 7	<0.000 1
7	612.702 6	<0.000 1	237.472 4	<0.000 1
8	650.545 3	<0.000 1	237.474 1	<0.000 1
9	691.057 5	<0.000 1	238.360 9	<0.000 1
10	748.420 0	<0.000 1	246.576 4	<0.000 1
11	773.679 6	<0.000 1	246.602 7	<0.000 1
12	807.689 5	<0.000 1	247.046 6	<0.000 1

Q 检验和LM 检验12阶延迟都显示该序列显著方差非齐，这说明残差平方序列中存在长期的相关关系。这种情况下，通常可以用高阶 ARCH 模型或者低阶GARCH 模型提取残差平方序列中蕴涵的相关关系。

GARCH模型的参数估计

- 条件最小二乘估计方法
  - GARCH(p,q)模型的误差平方和为

$$Q=\sum_{t=p+q+1}^{T-1}\left(\varepsilon_t^2-h_t\right)^2=\sum_{t=p+q+1}^{T-1}\left(\varepsilon_t^2-\lambda_0-\sum_{i=1}^{q}\lambda_i\varepsilon_{t-i}^2-\sum_{i=1}^{p}\eta_i h_{t-i}\right)^2$$

  - 使得Q 达到最小的参数值即该GARCH模型的条件最小二乘估计值
- 极大似然估计方法
  - GARCH(p,q)模型的对数似然函数为

$$\ln L(\omega,\lambda_1,\cdots,\lambda_q,\eta_1,\cdots,\eta_p)=-\frac{n}{2}\ln(2\pi)-\frac{1}{2}\sum_{t=1}^{n}\left[\sum_{t=1}^{n}(\lambda_0+\sum_{i=1}^{q}\lambda_i\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i h_{t-i})+\frac{\varepsilon_t^2}{2(\lambda_0+\sum_{i=1}^{q}\lambda_i\varepsilon_{t-i}^2+\sum_{i=1}^{p}\eta_i h_{t-i})}\right]$$

  - 使得似然函数最大的参数值即为该GARCH模型的极大似然估计值

例7-2续

- 对2013年1月4日至2017年8月25日上证指数每日收盘价序列拟合GARCH模型，并估计模型的未知参数。
  - 要拟合GARCH模型首先需要对模型定阶。Q检验和LM检验结果显示残差平方序列具有长期相关。这种情况下，通常是尝试拟合高阶ARCH模型或低阶GARCH模型。
  - 拟合模型一：ARCH（4）模型。基于极大似然估计得到的拟合模型为

$$\begin{cases} x_t = x_{t-1} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t \quad , e_t \sim N(0,1) \\ h_t = 385.2571 + 0.0431\varepsilon_{t-1}^2 + 0.4711\varepsilon_{t-2}^2 + 0.2687\varepsilon_{t-3}^2 + 0.2328\varepsilon_{t-4}^2 \end{cases}$$

  - 拟合模型二：GARCH(1,1)模型。基于条件最小二乘估计得到的拟合模型为

$$\begin{cases} x_t = x_{t-1} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t \quad , e_t \sim N(0,1) \\ h_t = 5.5933 + 0.0663\varepsilon_{t-1}^2 + 0.9309 h_{t-1} \end{cases}$$

拟合检验

- GARCH模型拟合出来之后，我们需要对它进行拟合检验。检验内容主要包括如下三个方面
  - 参数显著性检验
    - 参数显著性检验和ARIMA模型的参数显著性检验一样，构造t分布检验统计量
    - 在显著性水平取为 α 时，如果t统计量的P值小于α，认为该参数显著非零。反之，参数不显著非零，可以删除该参数。
  - 模型显著性检验
  - 分布检验

模型显著性检验

- 残差序列标准化，残差平方序列标准化

$$e_t = \frac{\varepsilon_t}{\sqrt{\hat{h}_t}} \quad , \quad e_t^2 = \frac{\varepsilon_t^2}{\hat{h}_t}$$

- 模型显著成立的要求：
  - 均值模型中相关信息提取充分——标准化残差序列为白噪声序列
  - 方差模型中相关信息提取充分——标准化残差平方序列为白噪声序列

分布检验

- 图检验方法
  - 残差序列的QQ图
  - 残差序列的直方图
- 统计检验方法
  - 正态分布假定下，可以使用JB检验统计量
  - 该统计量的构造思想是，借助正态分布的偏态系数和峰态系数构造出一个服从自由度为2的卡方分布统计量

$$H_0: \frac{\varepsilon_t}{\sqrt{h_t}} \sim N(0,1) \leftrightarrow H_1: \frac{\varepsilon_t}{\sqrt{h_t}} \overset{\text{not}}{\sim} N(0,1)$$

$$JB = \frac{T}{6} b_1^2 + \frac{T}{24}\left(b_2^2 - 3\right)^2 \sim \chi^2(2)$$

例7-2续

- 对2013年1月4日至2017年8月25日上证指数每日收盘价序列进行拟合检验和模型优化。
  - 模型比较：GARCH(1,1)模型的AIC和BIC信息量都比ARCH(4)模型小，所以这两个拟合模型进行比较，GARCH(1,1)模型相对更优

模型	AIC	BIC
ARCH(4)	11307.6996	11332.8406
GARCH(1,1)	11108.4279	11123.5125

  - GARCH(1,1)模型拟合检验

（1）参数显著性检验：所有参数均显著非零

参数	t 统计量	P 值
λ_0	3. 09	0. 002
λ_1	10. 47	$<$0. 000 1
η_1	181. 10	$<$0. 000 1

例7-2续

- 模型显著性检验
  - 标准化残差序列白噪声检验

至滞后	卡方	自由度	$Pr>$卡方	自相关					
6	11.36	6	0.078 0	0.059	0.001	0.010	0.076	−0.026	−0.009
12	24.01	12	0.020 3	0.029	0.057	0.064	0.030	−0.022	−0.040
18	27.13	18	0.076 6	0.018	−0.024	0.024	0.019	0.017	0.026
24	31.32	24	0.144 7	0.023	0.034	−0.024	−0.031	−0.014	0.015

  - 标准化残差平方序列白噪声检验

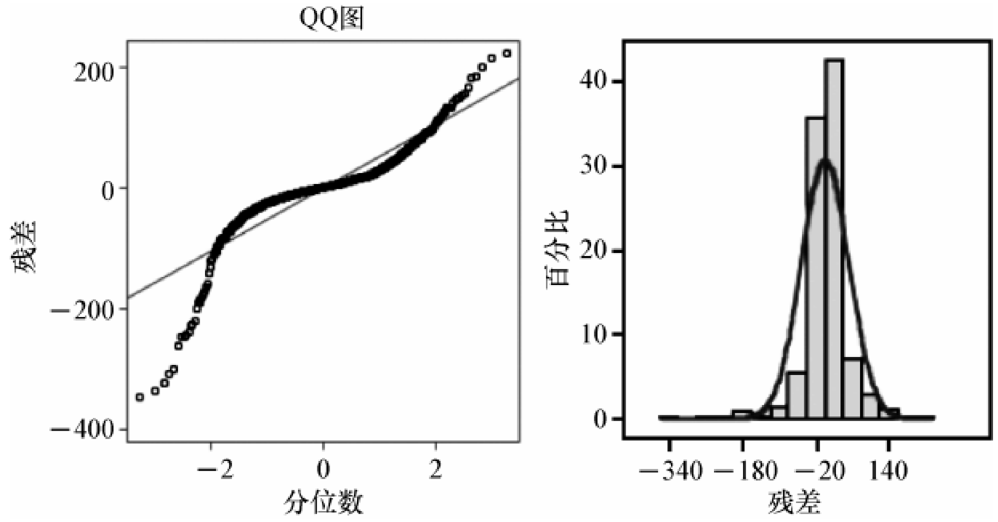
至滞后	卡方	自由度	$Pr>$卡方	自相关					
6	8.42	6	0.208 6	−0.040	0.049	0.036	0.007	0.008	−0.045
12	11.61	12	0.477 1	0.034	−0.032	−0.002	0.016	−0.008	−0.017
18	17.60	18	0.482 5	0.006	0.017	−0.004	0.046	0.039	−0.035
24	20.57	24	0.663 9	0.005	0.037	0.012	−0.008	−0.020	−0.023

- 均值模型对水平信息提取不充分
- 需要进一步考虑是水平定阶不准确造成的，还是分布假定不合适造成的

- 条件异方差信息提取充分

正态分布检验

- 图检验



图检验和JB检验结果都显示：

上证指数序列不服从正态分布。
如果对现有拟合精度不满意，应该选择比正态分布更尖峰的假定分布。

- JB检验

JB 检验统计量	自由度	P 值
230.3485	2	<0.0001

模型预测

- 序列预测值方差等于

$$Var(\hat{x}_{t+k}) = Var(\varepsilon_{t+k}) + \psi_1^2 Var(\varepsilon_{t+k}) + \cdots + \psi_{k-1}^2 Var(\varepsilon_{t+1})$$

- 条件异方差场合，序列预测方差等于

$$Var(\hat{x}_{t+k}) = \hat{h}_{t+k} + \psi_1^2 \hat{h}_{t+k-1} + \cdots + \psi_{k-1}^2 \hat{h}_{t+1}$$

- 正态分布假定下，预测值95%置信区间

$$\left(\hat{x}_{t+k} - 2\sqrt{Var(\hat{x}_{t+k})}, \hat{x}_{t+k} + 2\sqrt{Var(\hat{x}_{t+k})}\right), \forall k \geq 1$$

例7-2续

- 基于2013年1月4日至2017年8月25日上证指数每日收盘价序列，分别在方差齐性假定和条件异方差假定下，求该序列未来5期的95%置信区间
- 方差齐性假定下的预测值与95%置信区间

k	预测值	标准差	95%置信下限	95%置信上限
1	3 331.522	51.988	3 229.627 4	3 433.416 6
2	3 331.522	73.522	3 187.421 2	3 475.622 7
3	3 331.522	90.045	3 155.035 3	3 508.008 6
4	3 331.522	103.976	3 127.732 8	3 535.311 2
5	3 331.522	116.248	3 103.678 7	3 559.365 2

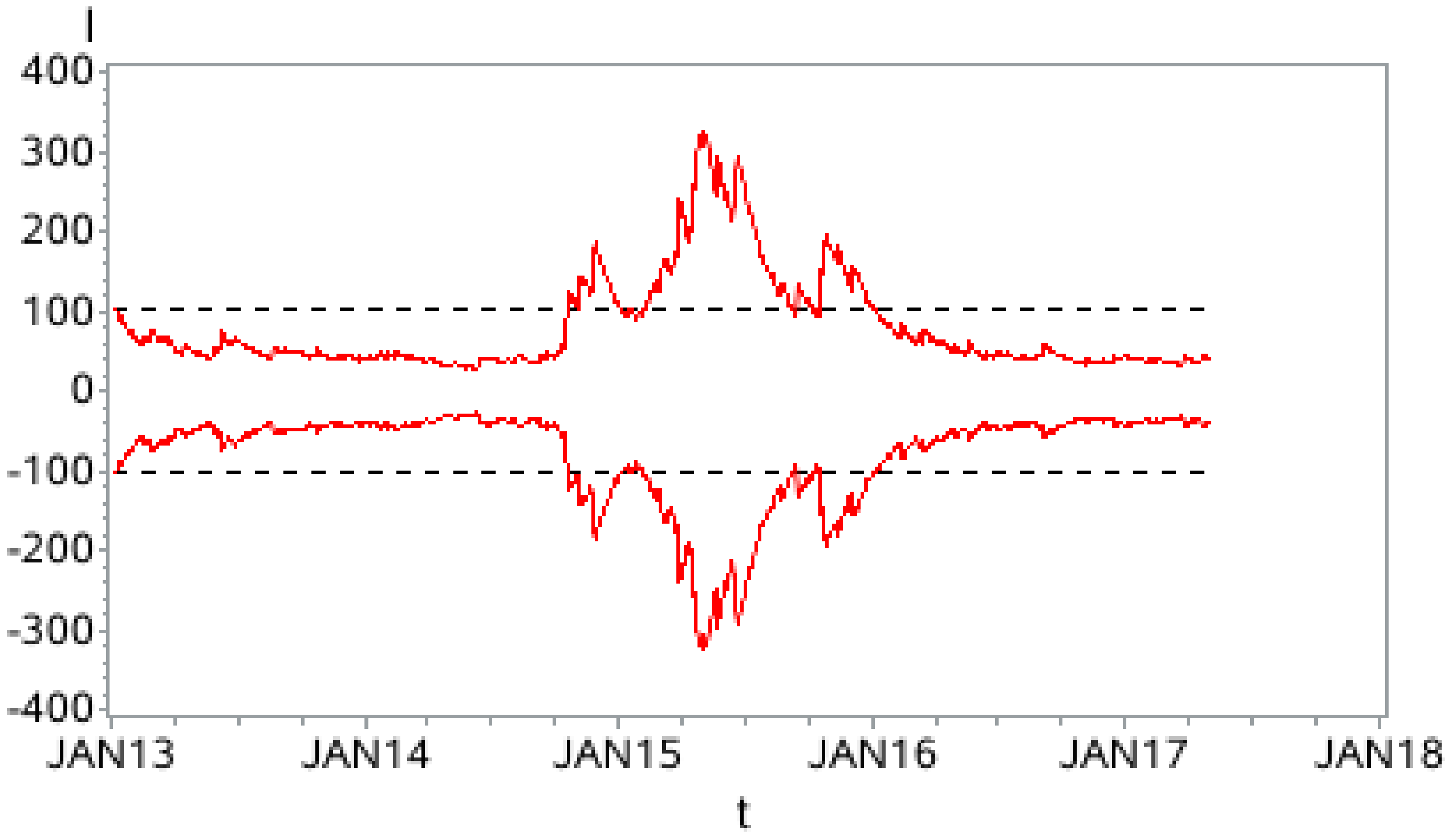
- 方差非齐假定下的预测值与95%置信区间

k	预测值	标准差	95%置信下限	95%置信上限
1	3 331.522	19.58	3 292.362	3 370.682
2	3 331.522	27.94	3 275.642	3 387.402
3	3 331.522	34.31	3 262.902	3 400.142
4	3 331.522	39.79	3 251.942	3 411.102
5	3 331.522	44.55	3 242.422	3 420.622

- 可以查到未来五个工作日上证指数的真实收盘价为
- 3362.65
  3365.23
  3363.63
  3360.81
  3367.12
- 两种假定下得到的置信区间都包含了序列真实值。但条件异方差假定下得到的置信区间范围更窄，说明它的精度更高。

例7-2续

两种方差假定下置信区间比较图



从图中我们可以清楚地看到，当序列大幅波动时条件异方差置信区间更宽，序列小幅波动时条件异方差置信区间更窄。这说明条件异方差模型对序列波动风险的拟合和预测通常更准确。

本章内容

GARCH模型的不足

（1）它对参数的约束条件非常严格，参数的约束条件一定程度上限制了GARCH模型的适用面

- 方差非负的要求，导致了参数非负的约束条件

$$\omega > 0, \eta_i \geq 0, \lambda_j \geq 0$$

- 无条件方差存在且方差非负的要求，导致了参数和小于1的约束条件

$$\sum_{i=1}^{p} \eta_i + \sum_{j=1}^{q} \lambda_j < 1$$

GARCH模型的不足

（2）它对正负扰动的反应是对称的，这与实际情况不符。

- 扰动项是真实值与预测值之差。如果扰动项为正，说明真实值比预测值大，对于投资者而言就是获得超预期收益。如果扰动项为负，说明真实值比预测值小，对于投资者而言就是出现了超预期的亏损。
- 以模型ARCH（1）为例，$\varepsilon_t^2=\phi\varepsilon_{t-1}^2$， 理论上无论 ε_{t-1}是正是负，它对下一期的影响系数都是 ϕ 。这意味着无论上一期的投资是收益还是亏损，对投资人下一期的投资行为的影响是一样的，这与实际情况不符。
- 大量的实践经验显示，投资人在面对收益和亏损时的反应不是对称的。出现收益时，通常反应比较慢；出现亏损时，通常反应比较快。忽视这种信息的不对称性，有时会影响预测的精度。

GARCH 衍生模型

- 基于GARCH模型的这些不完善的地方，产生了很多GARCH衍生模型
- 我们在此介绍使用最多、发展最为成熟的三个衍生模型：
  - EGARCH
  - IGARCH
  - GARCH-M

EGARCH模型

- Nelson于1991年提出Exponential GARCH模型，我们称之为指数GARCH模型，简记为EGARCH。
- EGARCH(p,q)模型完整结构如下

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \cdots) + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t \\ \ln h_t = \omega + \sum_{i=1}^{p} \eta_i \ln h_{t-i} + \sum_{j=1}^{q} \lambda_j g(e_{t-j}) \\ g(e_t) = \theta e_t + \gamma [\, |e_t| - E|e_t| \,] \end{cases}$$

EGARCH模型的改进

- 第一个改进：响应变量由条件方差 h_t 变成了条件方差的对数 $\ln(h_t)$ 。 $\ln(h_t)$ 可以是正数，也可以是负数。这导致参数不需要满足方差非负假定。因此，EGARCH模型的第一个改进是放松了对GARCH模型的参数约束。
- 第二个改进：自变量由误差平方 ε_t^2 变成了加权扰动函数 $g(e_t)$ 。通过特殊的函数构造，$g(e_t)$ 能对正负扰动进行非对称处理。

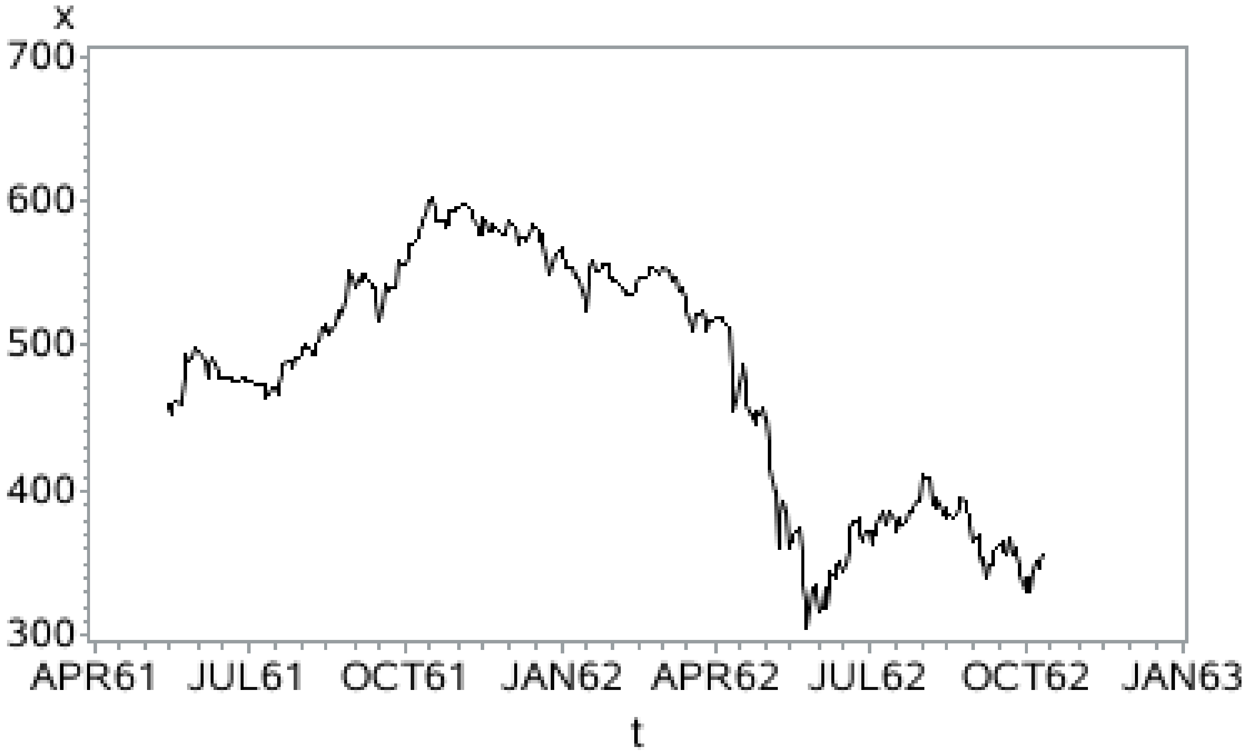
$$
\begin{aligned}
g(e_t) &= \theta e_t + \gamma[|e_t| - E|e_t|] \\
&= \begin{cases} (\theta+\gamma)e_t - \gamma E|e_t|, & e_t > 0 \\ (\theta-\gamma)e_t - \gamma E|e_t|, & e_t < 0 \end{cases}
\end{aligned}
$$

式中，$E[g(e_t)] = 0$；$e_t \sim N(0,1)$；$E|e_t| = \sqrt{2/\pi}$。通常取 $\gamma=1$，则 $g(e_t)$ 函数简写为：

$$
g(e_t) = \begin{cases} (\theta+1)e_t - \sqrt{2/\pi}, & e_t > 0 \\ (\theta-1)e_t - \sqrt{2/\pi}, & e_t < 0 \end{cases}
$$

例7-3

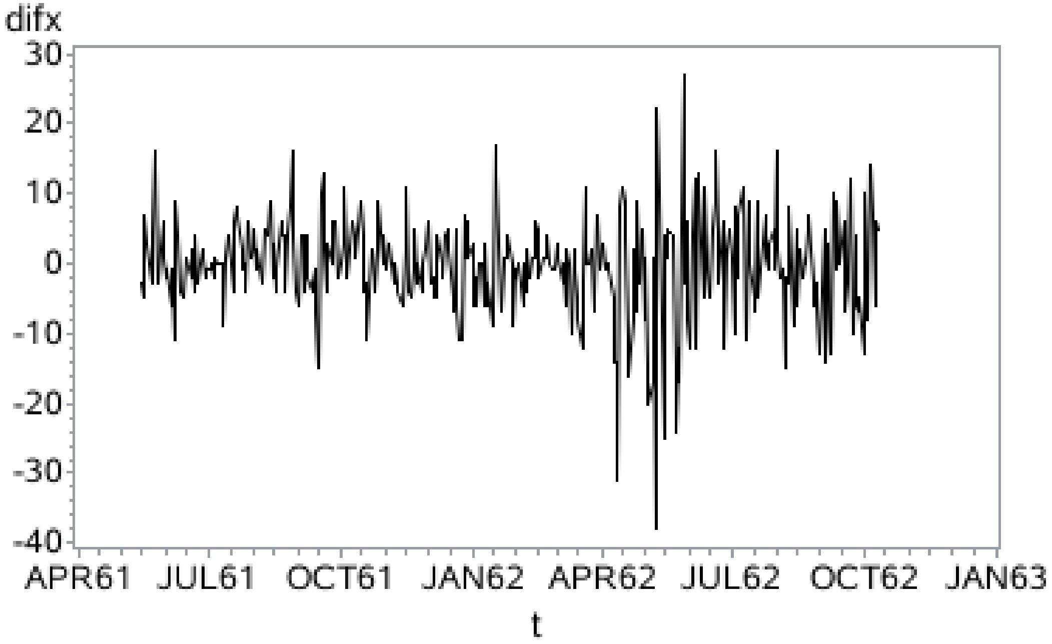
- 拟合1961年 5 月17日至1962年11月 2 日IBM股票每日收盘价序列。



序列差分平稳

- 对IBM收盘价序列进行1阶差分，检验结果显示，该序列1阶差分后平稳

1阶差分后残差图



1阶差分后残差序列PP检验结果

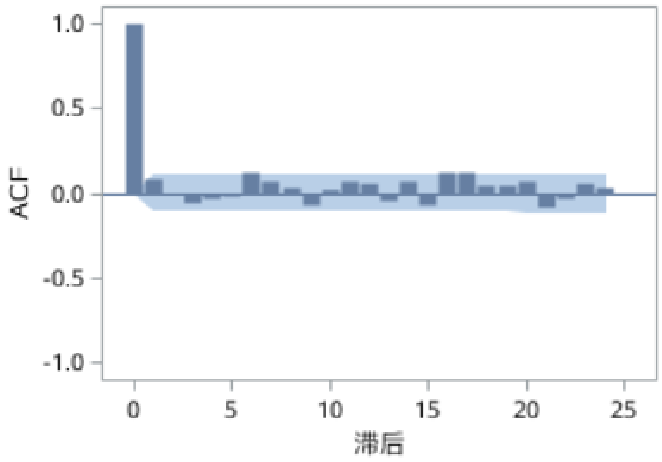
类型	延迟阶数	τ 统计量的值	$Pr < \tau$
类型一	0	-17.52	<.0001
	1	-17.52	<.0001
	2	-17.52	<.0001
类型二	0	-17.52	<.0001
	1	-17.52	<.0001
	2	-17.52	<.0001
类型三	0	-17.64	<.0001
	1	-17.64	<.0001
	2	-17.64	<.0001

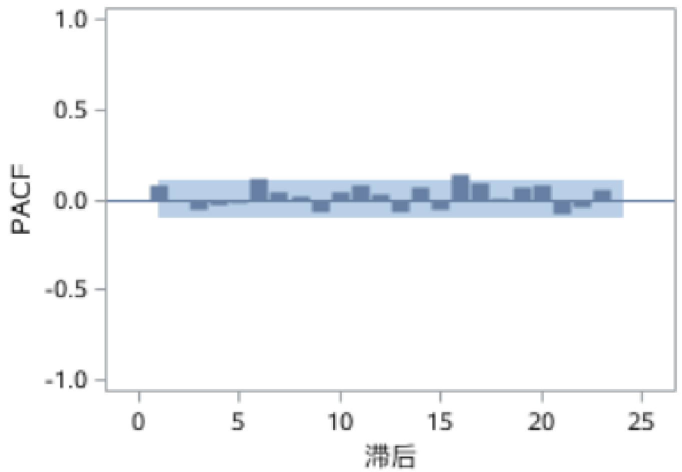
白噪声检验

- 差分后序列为白噪声序列

延迟阶数	纯随机性检验	
	LB 检验统计量的值	P 值
6	9.98	0.1256
12	17.42	0.1344

- 差分后序列的自相关图和偏自相关图也显示出白噪声属性





条件异方差检验

- Q统计量和LM检验都显示，残差序列具有长期相关的条件异方差属性。

阶数	Q	Pr > Q	LM	Pr > LM
1	25.0633	<.0001	24.9351	<.0001
2	34.2547	<.0001	28.1337	<.0001
3	57.4705	<.0001	42.1683	<.0001
4	82.0689	<.0001	50.6202	<.0001
5	82.0766	<.0001	56.8440	<.0001
6	86.8290	<.0001	58.3769	<.0001
7	88.1800	<.0001	59.3373	<.0001
8	89.1418	<.0001	59.4981	<.0001
9	103.8702	<.0001	74.4553	<.0001
10	118.8154	<.0001	76.8409	<.0001
11	121.0637	<.0001	76.8427	<.0001
12	129.8700	<.0001	77.3868	<.0001

条件异方差模型拟合一

- 拟合模型一：GARCH(1,1)

$$\begin{cases} x_t = x_{t-1} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t \quad , e_t \sim N(0, 52.67) \\ h_t = 4.5142 + 0.6563 h_{t-1} + 0.2704 \varepsilon_{t-1}^2 \end{cases}$$

- 该拟合模型的AIC=2504.15，BIC=2508.06

条件异方差模型拟合二

- 拟合模型二：EGARCH(1,1)
- 拟合理由：正负扰动项不均衡

	N	ε_t^2 的均值	ε_t^2 的标准差
$\varepsilon_t>0$	166	48.644 578 3	85.571 52
$\varepsilon_t<0$	169	66.792 899 4	156.303 005 3

- 拟合EGARCH(1,1)模型

$$\ln h_t = 0.3345 + 0.911\ln h_{t-1} + 0.2791 g\left(e_{t\text{-}1}\right)$$

$$g\left(e_t\right)=\begin{cases}(-0.515+1)e_t-\sqrt{2/\pi} & ,e_t>0\\(-0.515-1)e_t-\sqrt{2/\pi} & ,e_t<0\end{cases}$$

- 该拟合模型的AIC=2409.44，BIC=2428.98。本例EGARCH模型优于GARCH模型。

IGARCH模型

- Nelson于1990年提出integrated GARCH模型，直译为集成GARCH模型，也称为方差无穷GARCH模型，简记为IGARCH(p,q)。
- IGARCH(p,q)模型的完整结构如下

$$
\begin{cases}
x_t = f(t, x_{t-1}, x_{t-2}, \cdots) + \varepsilon_t \\
\varepsilon_t = \sqrt{h_t} e_t \quad , \quad e_t \overset{i.i.d}{\sim} N(0,1) \\
h_t = \lambda_0 + \sum_{i=1}^{p} \eta_i h_{t-i} + \sum_{j=1}^{q} \lambda_j \varepsilon_{t-j}^2 \\
\sum_{i=1}^{p} \eta_i + \sum_{j=1}^{q} \lambda_j = 1
\end{cases}
$$

IGARCH模型的改进

- 和GARCH模型相比，IGARCH模型增加了一个参数约束条件

$$\sum_{i=1}^{p} \eta_i + \sum_{j=1}^{q} \lambda_j = 1$$

- 因为残差序列的无条件方差为

$$Var(\varepsilon_t) = \frac{\omega}{1 - \left(\sum_{i=1}^{p} \eta_i + \sum_{j=1}^{q} \lambda_j \right)}$$

- 增加的这个约束条件导致残差序列的方差无界（允许方差无穷大）。

IGARCH模型的使用环境

- IGARCH模型适合于描述具有单位根特征的条件异方差。从理论角度来说，IGARCH 现象可能是由波动率带有常数漂移项引起的。
- 以IGARCH(1,1）模型为例，条件方差预测值为

$$
\begin{aligned}
\sigma_t^2(1) &= \omega + \sigma_t^2 \\
\sigma_t^2(2) &= \omega + \sigma_t^2(1) = 2\omega + \sigma_t^2 \\
&\vdots \\
\sigma_t^2(h) &= \omega + \sigma_t^2(h-1) = h\omega + \sigma_t^2
\end{aligned}
$$

例7-3续

- 对1961年5月17日至1962年11月2日IBM股票每日收盘价序列拟合IGARCH模型。

$$\begin{cases} x_t = x_{t-1} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t, e_t \sim N(0, 52.66) \\ h_t = 3.1061 + 0.6784 h_{t-1} + 0.3216 \varepsilon_{t-1}^2 \end{cases}$$

- 该拟合模型的AIC=2419.30，BIC=2431.01。

GARCH-M模型产生背景

- 金融中，风险厌恶型的投资者会要求资产的收益率与它的波动性相匹配。

风险投资期望收益=无风险收益+风险溢价

  - 无风险收益（risk-free return）：即为资金的时间价值，是投资者从事风险极小投资（诸如国库券，货币市场工具或银行存款）所获得的收益。
  - 超额收益（excess return）：任何特定时期风险资产同无风险资产收益之差称为超额收益。
  - 风险溢价（risk premium）：超额收益的期望称为风险溢价。风险溢价指投资者因承担风险而获得的额外报酬。风险越大，风险溢价则越高。

- Engle，Lilien，Robins在1987年将风险溢价思想引入GARCH模型，允许序列的均值依赖于它的波动性，由此提出GARCH-Mean模型，简记为GARCH-M。

GARCH-M模型结构

- 模型构造思想
  - 序列均值会受到序列条件方差的影响，这时可以把条件标准差作为附加回归因子建模

- 模型结构

$$\begin{cases} x_t = f(t, x_{t-1}, x_{t-2}, \cdots) + \delta\sqrt{h_t} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t}e_t \\ h_t = \omega + \sum_{i=1}^{p} \eta_i h_{t-i} + \sum_{j=1}^{q} \lambda_j \varepsilon_{t-j}^2 \end{cases}$$

例7-3续

- 对1961年5月17日至1962年11月2日IBM股票每日收盘价序列拟合GARCH-M模型

$$\begin{cases} x_t = x_{t-1} + 0.1327\sqrt{h_t} + \varepsilon_t \\ \varepsilon_t = \sqrt{h_t} e_t \quad , e_t \sim N(0, 52.90) \\ h_t = 5.0928 + 0.6247 h_{t-1} + 0.2934 \varepsilon_{t-1}^2 \end{cases}$$

- 该拟合模型的AIC=2420.15，BIC=2439.69。

07

THANKS