第 22 章　自回归条件异方差模型

22.1　条件异方差模型的例子

Engle (1982)指出，时间序列数据也常存在一种特殊的异方差，即“自回归条件异方差”(Autoregressive Conditional Heteroskedasticity，简记 ARCH)。

Bollerslev (1986)对 ARCH 进行了推广，称为“Generalized ARCH”，简记 GARCH。

考察美国道琼斯股指在 1953—1990 年期间日收益率的波动。

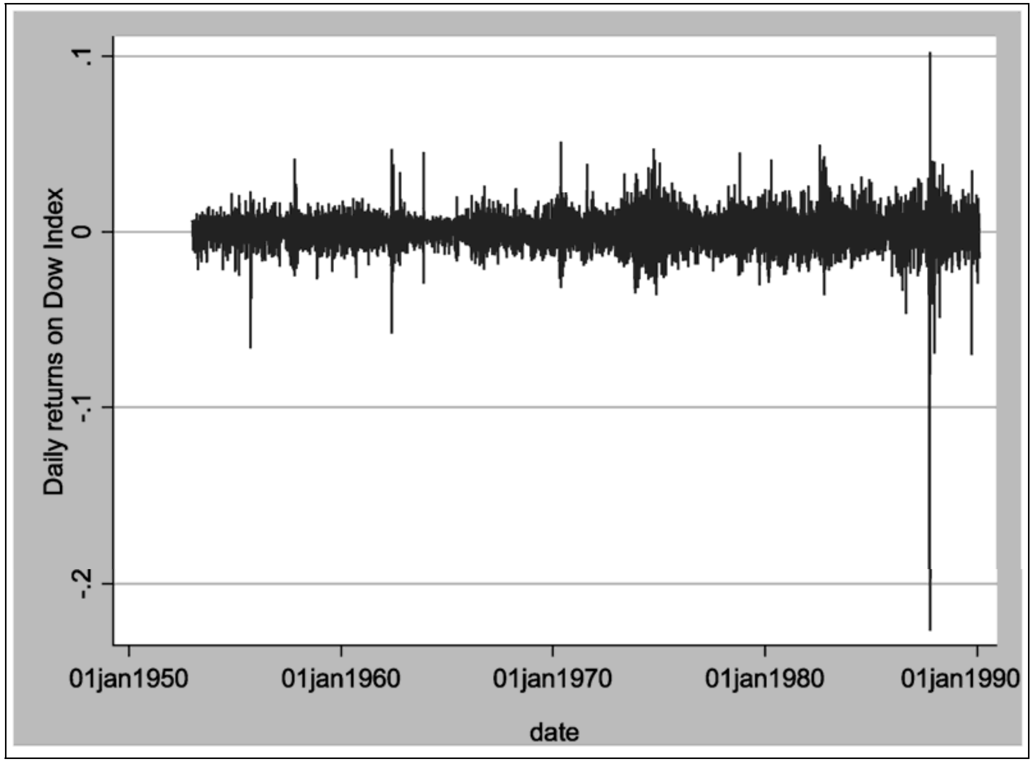

图 22.1　美国道琼斯股指 1953—1990 的日收益率

从图 22.1 可知，方差大的观测值似乎集聚在一起，而方差小的观测值似乎也集聚在一起，称为“波动性集聚”(volatility clustering)。

之前，由于缺乏更好的度量，一般假设时间序列的方差恒定。

由于 ARCH 模型考虑了方差的波动性，故可更好地预测方差，在金融领域有重要应用价值。

22.2 ARCH 模型的性质

考虑线性回归模型：

$$y_t = \boldsymbol{x}_t' \boldsymbol{\beta} + \varepsilon_t$$

记扰动项ε_t的条件方差为$\sigma_t^2 \equiv \mathrm{Var}(\varepsilon_t \mid \varepsilon_{t-1}, \cdots)$，可随时间而变。

受波动性集聚启发，假设

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

这就是“ARCH(1)扰动项”。更一般地，假设

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2$$

这就是“ARCH(p)扰动项”。

以 ARCH(1)为例考察 ARCH 扰动项的性质。

假设扰动项ε_t的生成过程为

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$$

其中，v_t为白噪声，方差标准化为 1，即$\mathrm{Var}(v_t) = \mathrm{E}(v_t^2) = 1$。

假定v_t与ε_{t-1}相互独立，且$\alpha_0 > 0$，$0 < \alpha_1 < 1$ (为了保证σ_t^2为正，且$\{\varepsilon_t\}$为平稳过程)。

由于v_t与ε_{t-1}相互独立，ε_t的条件期望为

$$\begin{aligned}\mathrm{E}(\varepsilon_t \mid \varepsilon_{t-1}) &= \mathrm{E}\left\{ v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \,\middle|\, \varepsilon_{t-1} \right\} \\ &= \underbrace{\mathrm{E}(v_t)}_{=0} \cdot \mathrm{E}\left\{ \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \,\middle|\, \varepsilon_{t-1} \right\} = 0\end{aligned}$$

类似地，ε_t的无条件期望为

$$\mathrm{E}(\varepsilon_t) = \mathrm{E}\left\{ v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \right\} = \underbrace{\mathrm{E}(v_t)}_{=0} \cdot \mathrm{E}\left\{ \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \right\} = 0$$

根据v_t与ε_{t-1}独立性，ε_t的条件方差为

$$\mathrm{Var}(\varepsilon_t \mid \varepsilon_{t-1}) = \mathrm{E}(\varepsilon_t^2 \mid \varepsilon_{t-1}) = \underbrace{\mathrm{E}(v_t^2)}_{=1} \cdot \mathrm{E}(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \mid \varepsilon_{t-1})$$

$$= \alpha_0 + \alpha_1 \mathrm{E}(\varepsilon_{t-1}^2 \mid \varepsilon_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

这就是 ARCH(1)的定义式“$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$”。

考察ε_t的无条件方差：

$$\mathrm{Var}(\varepsilon_t) = \mathrm{E}(\varepsilon_t^2) = \mathrm{E}\left[v_t^2 (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) \right]$$

$$= \underbrace{\mathrm{E}(v_t^2)}_{=1} \cdot \mathrm{E}(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) = \alpha_0 + \alpha_1 \mathrm{E}(\varepsilon_{t-1}^2)$$

对于差分方程“$\mathrm{E}(\varepsilon_t^2)=\alpha_0+\alpha_1\mathrm{E}(\varepsilon_{t-1}^2)$”，由于$0<\alpha_1<1$，故该差分方程有稳定解。

令$\mathrm{E}(\varepsilon_t^2)=\mathrm{E}(\varepsilon_{t-1}^2)$，可得$\mathrm{E}(\varepsilon_t^2)=\dfrac{\alpha_0}{1-\alpha_1}$，故 ARCH 扰动项的无条件方差为常数。

考察ε_t与$\varepsilon_{t-i}\ (i\neq 0)$的序列相关：

$$\mathrm{E}(\varepsilon_t\varepsilon_{t-i})=\mathrm{E}\left\{v_t v_{t-i}\sqrt{(\alpha_0+\alpha_1\varepsilon_{t-1}^2)(\alpha_0+\alpha_1\varepsilon_{t-i}^2)}\right\}$$

$$=\underbrace{\mathrm{E}(v_t v_{t-i})}_{=0}\cdot\mathrm{E}\left\{\sqrt{(\alpha_0+\alpha_1\varepsilon_{t-1}^2)(\alpha_0+\alpha_1\varepsilon_{t-i}^2)}\right\}=0$$

扰动项$\{\varepsilon_t\}$满足古典模型关于“同方差”与“无自相关”的假定，故高斯-马尔可夫定理成立，OLS 是 BLUE。

虽然$\{\varepsilon_t\}$存在条件异方差，却是白噪声！

但 OLS 忽略了条件异方差这一重要信息，可找到更优的非线性估计，即 MLE。

22.3 ARCH 模型的 MLE 估计

对于 ARCH(1)，为保证条件方差$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$始终非负，须限制参数$\alpha_0, \alpha_1$均为正数。

如果$\alpha_0 < 0$或$\alpha_1 < 0$，则可能出现“$\sigma_t^2 < 0$”的情形，参见图 22.2。另外，$\alpha_1 < 1$是为了保证$\{\varepsilon_t\}$为平稳过程。

如果$\alpha_1 > 1$，则$\text{Var}(\varepsilon_t)$将随时间而增大，不是平稳过程。

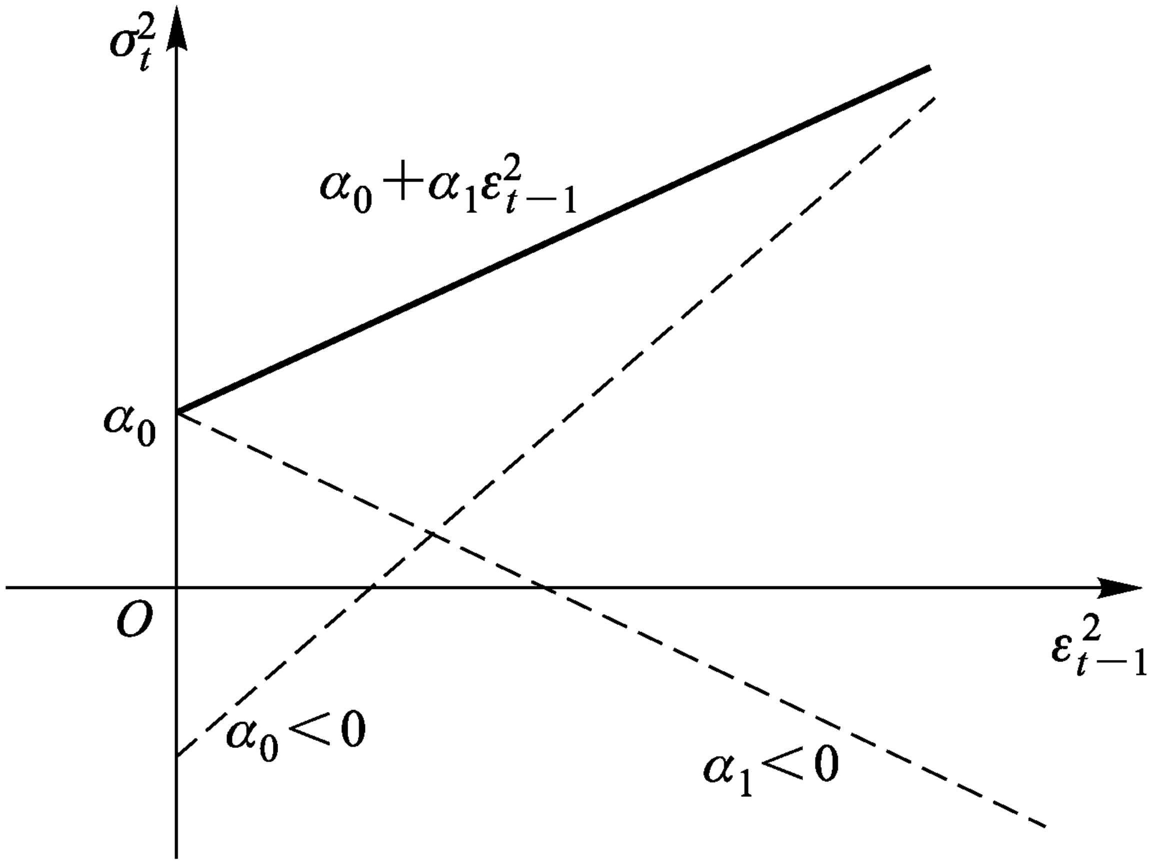

图 22.2　对 ARCH(1)参数的正约束

假设样本容量为 T。在 ARCH(1)模型中，$\{\varepsilon_t\}$并非 iid，因为相邻扰动项通过公式“$\varepsilon_t = v_t\sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}^2}$”联系在一起。

由于ε_t仅依赖于ε_{t-1}，而不依赖于$\{\varepsilon_{t-2}, \varepsilon_{t-3}, \cdots\}$，故可将$\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T\}$的联合密度函数分解如下：

$$\begin{aligned} f(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T) &= f(\varepsilon_1) f(\varepsilon_2 \mid \varepsilon_1) f(\varepsilon_3 \mid \varepsilon_2, \varepsilon_1) \cdots f(\varepsilon_T \mid \varepsilon_{T-1}, \cdots) \\ &= f(\varepsilon_1) f(\varepsilon_2 \mid \varepsilon_1) f(\varepsilon_3 \mid \varepsilon_2) \cdots f(\varepsilon_T \mid \varepsilon_{T-1}) \end{aligned}$$

由于无条件密度函数$f(\varepsilon_1)$不易计算，考虑给定ε_1情况下的条件MLE)。

假设$\varepsilon_t \sim N(0, \sigma_t^2)$，而$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$，可得似然函数：

$$L = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)}} \exp\left\{ -\frac{\varepsilon_t^2}{2(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)} \right\}$$

$$\ln L = -\frac{T-1}{2} \ln 2\pi - \frac{1}{2} \sum_{t=2}^{T} \ln(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) - \frac{1}{2} \sum_{t=2}^{T} \frac{\varepsilon_t^2}{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$$

将“$\varepsilon_t = y_t - \boldsymbol{x}_t' \boldsymbol{\beta}$”代入上式，则$\ln L$成为参数$(\alpha_0, \alpha_1, \beta_0, \beta_1)$的函数。

对 ARCH 模型进行 MLE 估计时，对原方程$(y_t = \boldsymbol{x}_t' \boldsymbol{\beta} + \varepsilon_t)$与条件方差方程$(\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2)$同时进行估计。

如果要估计 ARCH(p)，则将$\left\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p\right\}$(即前 p 个观测值)视为给定，然后使用条件 MLE。

即使扰动项不服从正态分布，作为 QMLE，仍可能是一致的。

22.4 GARCH 模型

在 ARCH(p)模型中，如 p 很大，要估计很多参数，损失样本容量。

Bollerslev (1986)提出 GARCH，使待估参数减少，对条件方差的预测更准确。

在 ARCH 模型的基础上，再加上σ_t^2的自回归部分。

GARCH(p, q)的模型设定为：

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \gamma_1 \sigma_{t-1}^2 + \cdots + \gamma_p \sigma_{t-p}^2$$

其中，p 为σ_t^2的自回归阶数，而 q 为ε_t^2的滞后阶数。

在 Stata 中，称ε_{t-i}^2为“ARCH 项”，而称σ_{t-i}^2为“GARCH 项”。

假定扰动项ε_t的生成过程为

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \gamma_1 \sigma_{t-1}^2 + \cdots + \gamma_p \sigma_{t-p}^2}$$

其中，v_t为白噪声。

最常见的 GARCH(1,1)：

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \sigma_{t-1}^2$$

其中，为了保证σ_t^2为正，$\alpha_0, \alpha_1, \gamma_1$均为正数。

GARCH(1,1)扰动项ε_t的生成过程为

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \sigma_{t-1}^2}$$

将方程两边平方，取(无条件)期望可得

$$
\begin{aligned}
\mathrm{E}(\varepsilon_t^2) &= \underbrace{\mathrm{E}(v_t^2)}_{=1} \cdot \mathrm{E}(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \sigma_{t-1}^2) \\
&= \alpha_0 + \alpha_1 \,\mathrm{E}(\varepsilon_{t-1}^2) + \gamma_1 \,\mathrm{E}(\sigma_{t-1}^2) \qquad\qquad (\sigma_{t-1}^2 \equiv \mathrm{E}(\varepsilon_{t-1}^2 \mid \varepsilon_{t-2})) \\
&= \alpha_0 + \alpha_1 \,\mathrm{E}(\varepsilon_{t-1}^2) + \gamma_1 \,\mathrm{E}\left[\mathrm{E}(\varepsilon_{t-1}^2 \mid \varepsilon_{t-2})\right] \quad (\text{迭代期望定律}) \\
&= \alpha_0 + \alpha_1 \,\mathrm{E}(\varepsilon_{t-1}^2) + \gamma_1 \,\mathrm{E}(\varepsilon_{t-1}^2) \\
&= \alpha_0 + (\alpha_1 + \gamma_1)\,\mathrm{E}(\varepsilon_{t-1}^2)
\end{aligned}
$$

为保证$\{\varepsilon_t\}$平稳，须要求$\alpha_1 + \gamma_1 < 1$。

为何使用 GARCH 模型能节省待估参数？因为σ_{t-1}^2中已经包含了$\left\{\varepsilon_{t-2}^2, \cdots, \varepsilon_{t-p-1}^2\right\}$的信息。

对 GARCH 模型可同样使用 MLE 估计。

22.5 何时使用 ARCH 或 GARCH 模型

初步方法可画时间序列图，看看是否存在“波动性集聚”。

严格的统计检验包括以下三种方法。

方法一 首先，用 OLS 估计原方程 $y_t = \boldsymbol{x}_t' \boldsymbol{\beta} + \varepsilon_t$，得到残差序列 $\{e_t\}$。其次，用 OLS 估计辅助回归，$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_p e_{t-p}^2 + error_t$，并检验原假设“$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$”(不存在条件异方差)。

Engle (1982)提出进行 LM 检验，其检验统计量为 $TR^2 \xrightarrow{d} \chi^2(p)$，其中 T 为样本容量，R^2 为辅助回归的可绝系数。

方法二 可以对残差平方序列$\left\{e_t^2\right\}$进行 Q 检验，检验其序列相关性。如果$\left\{e_t^2\right\}$存在自相关，则认为ε_t存在条件异方差。

方法三 最为直接的方法是，在估计 ARCH 或 GARCH 模型之后，看条件方差方程中的系数(即所有α与γ)是否显著。

22.6 ARCH 与 GARCH 模型的扩展

1. ARCH-M

金融理论认为，金融资产的风险越高，其期望收益率也应该越高。超出正常期望收益率的部分，称为“风险溢价”(risk premium)。

Engle, Lilien and Robins (1987)提出“ARCH-in-Mean 模型”(简记 ARCH-M)，假设金融资产的超额收益率满足以下方程：

$$y_t = \underbrace{\beta + \delta\sigma_t^2}_{\text{risk premium}} + \varepsilon_t$$

其中，y_t为“超额收益率”(excess return)，即超出无风险的国库券收益率的部分；ε_t为对超额收益率不可预见的冲击；而$(\beta + \delta\sigma_t^2)$为风险溢价，是条件方差$\sigma_t^2$的增函数，即$\delta > 0$。

假设ε_t服从 ARCH(p)过程，$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2$。

可使用 MLE 对原方程与条件方差方程同时进行估计。

2．TARCH

“坏消息”对资产价格波动性的影响可能大于好消息的影响。

Glosten, Jagannathan and Runkle (1993)提出了非对称(asymmetric)的“门限 GARCH”模型(Threshold GARCH，简记 TARCH)。

假设条件方差方程为

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \lambda_1 \underbrace{\varepsilon_{t-1}^2 \cdot \boldsymbol{1}(\varepsilon_{t-1} > 0)}_{tarch} + \beta_1 \sigma_{t-1}^2$$

Stata 称“$\varepsilon_{t-1}^2 \cdot \boldsymbol{1}(\varepsilon_{t-1} > 0)$”为 TARCH 项。

3．EGARCH

标准的 GARCH 模型对参数取值有限制，给 MLE 估计带来不便。

考虑以下对数形式的条件方差方程：

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 \underbrace{\left(\varepsilon_{t-1}/\sigma_{t-1}\right)}_{earch} + \lambda_1 \underbrace{\left|\varepsilon_{t-1}/\sigma_{t-1}\right|}_{earch\_a} + \beta_1 \underbrace{\ln \sigma_{t-1}^2}_{egarch}$$

其中，$\left(\varepsilon_{t-1}/\sigma_{t-1}\right)$为$\varepsilon_{t-1}$的标准化，　表示非对称效应。

$\left|\varepsilon_{t-1}/\sigma_{t-1}\right|$表示对称效应，Stata 称为“earch_a”项(a 表示“absolute value”，即绝对值)。Stata 称$\ln \sigma_t^2$为“egarch”项。

无论$\ln\sigma_t^2$取何值，都有$\sigma_t^2=\exp\left(\ln\sigma_t^2\right)>0$，故对所有参数都无限制。

由于σ_t^2为指数形式，故称为“指数 GARCH”(Exponential GARCH，简记 EGARCH)。

4．带 ARMA 的 GARCH

考虑如下线性回归模型：

$$y_t=\boldsymbol{x}_t'\boldsymbol{\beta}+u_t$$

其中，扰动项u_t为 ARMA(p, q)过程：

$$u_t=\sum\nolimits_{i=1}^{p}\rho_i u_{t-i}+\sum\nolimits_{j=1}^{q}\theta_j\varepsilon_{t-j}+\varepsilon_t$$

其中，ε_t为 GARCH(或 ARCH)扰动项。代入方程：

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \sum\nolimits_{i=1}^{p} \rho_i (y_{t-i} - \boldsymbol{x}_{t-i}'\boldsymbol{\beta}) + \sum\nolimits_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

这就是“带 ARMA 的 GARCH”(GARCH with ARMA terms)。

5．在条件方差方程中引入解释变量

例 为考察 911 恐怖袭击后波动性是否增大，引入虚拟变量

$$D_t = \begin{cases} 1, t \ge 2001/9/11 \\ 0, t < 2001/9/11 \end{cases}$$

然后考虑以下 GARCH(1, 1)模型：

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma D_t$$

6．使用非正态扰动项

如果被解释变量(比如，某些金融变量)的分布函数存在厚尾，则小概率事件比在正态分布情况下更容易发生。

可以选择让扰动项服从$t(k)$分布来估计 ARCH 或 GARCH 模型。

在进行 MLE 估计时，将 t 分布的自由度k也作为待估参数。