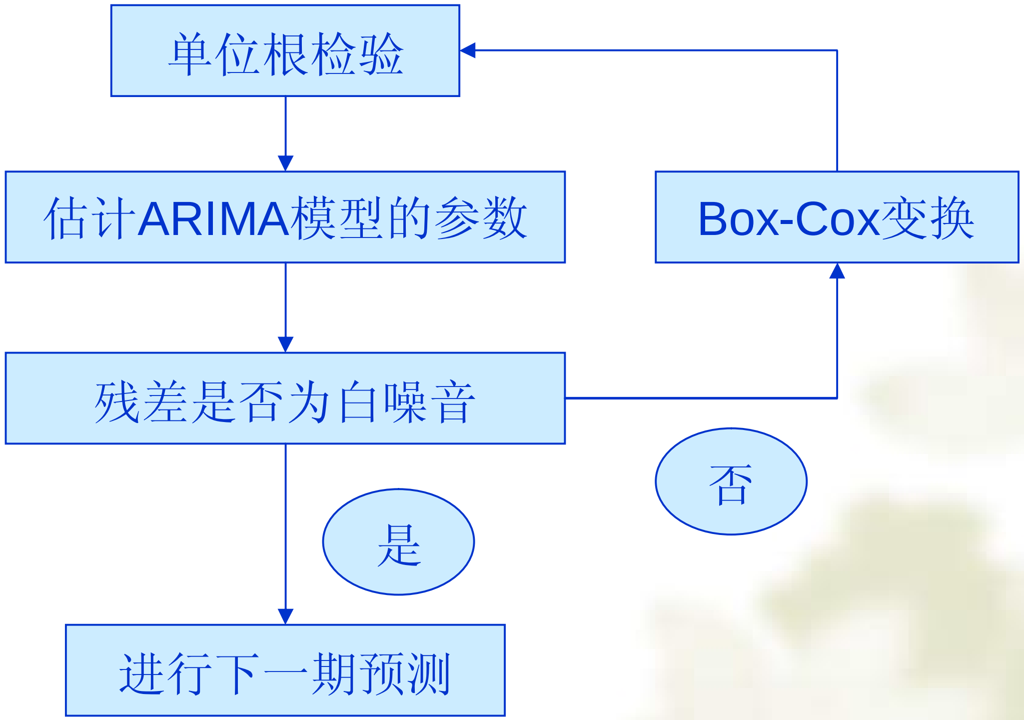
时间序列模型—
ARIMA与GARCH类模型

厦门大学财政系 王艺明

ARIMA模型

- 对于投资人而言，其最关心的莫过于能预先获知商品股价的变动趋势，获取最大的利润；而对于商品交易的管理当局而言，其亦期望能对未来市场趋于过热或过冷加以预测，而预先加以对策因应，规避市场的过度波动
- 因此传统的时间序列模型旨在提供一套有效的预测技术，以掌握近期内各金融商品股价的走向，提前因应景气转坏所受的冲击
- 其中最著名的之一便是Box & Jenkins 于1976年提出ARIMA模型，以往的文献曾将单变量ARIMA模型的预测效果与其它预测技术（如多元回归、GARCH、倒传递类神经网络模型、向量自回归模型及误差修正模型等）相比较，其结果显示ARIMA预测技术有一定的使用价值

ARIMA模型的研究步骤



ARMA模型的设定

- 我们假设经过差分后获得的时序y_t是平稳的，此时我们对其建立ARMA（p,q）模型
- ARMA模型假设时序y_t为其当前与前期的误差和随机项，以及它的前期值的线性函数，表示为
- $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \mu_t - \theta_1 \mu_{t-1} - \theta_2 \mu_{t-2} - \ldots - \theta_q \mu_{t-q}$
- 则称该时序y_t为自回归移动平均序列，该模型为（p,q）阶自回归移动平均模型，记作ARMA（p,q）。参数$\phi_1, \phi_2, \ldots, \phi_p$为自回归参数，$\theta_1, \theta_2, \ldots, \theta_q$为移动平均参数，是模型的待估参数

ARMA模型p,q阶数的确定

- 决定ARMA（p,q）模型中的p与q值阶数的方法之一为，应用样本的自相关图（ACF）判断AR（p）的阶数，应用偏自相关图（PACF）判断MA（q）的阶数
- 但该方法较主观

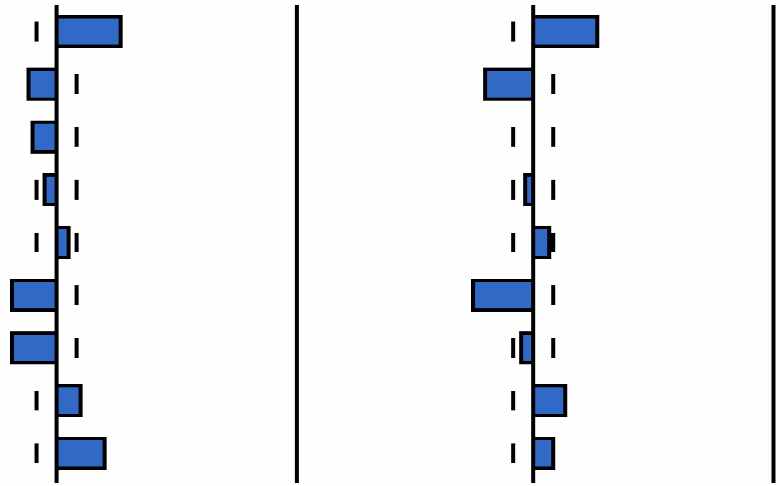
模型	ACF	PACF
AR(p)	呈指数或正弦函数图形	截断与k，当k>p
MA(q)	截断于q期之后	呈指数或正弦函数图形
ARMA(p,q)	呈指数与正弦函数混合的衰退消失	呈指数与正弦函数混合的衰退消失

美国月利率数据的拟合－ARMA

- 样本数据为第一讲中的数据
- 对月利率r进行单位根检验，无法拒绝其为单位根的零假设
- 因此应拟合ARIMA（p,1,q）模型
- 产生差分序列 new series dr=r-r(-1)
- 观察dr的自相关和偏自相关图形

美国月利率数据的拟合－ARMA

Date: 07/10/05 Time: 22:23
Sample: 1959:01 1996:02
Included observations: 445



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.272	0.272	33.024	0.000
		2	-0.116	-0.204	39.026	0.000
		3	-0.099	-0.007	43.466	0.000
		4	-0.044	-0.037	44.333	0.000
		5	0.053	0.064	45.599	0.000
		6	-0.183	-0.262	60.735	0.000
		7	-0.194	-0.050	77.902	0.000
		8	0.100	0.144	82.462	0.000
		9	0.202	0.089	101.00	0.000

美国月利率数据的拟合－ARMA

- 从PACF图形来看，似乎是截尾的，可以考虑AR(2)模型
- 输入 ls dr c ar(1) ar(2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.004472	0.027779	0.160969	0.8722
AR(1)	0.327216	0.046663	7.012350	0.0000
AR(2)	-0.204470	0.046664	-4.381785	0.0000

R-squared	0.112554	Mean dependent var	0.004555
Adjusted R-squared	0.108520	S.D. dependent var	0.543241
S.E. of regression	0.512918	Akaike info criterion	1.509348
Sum squared resid	115.7574	Schwarz criterion	1.537070
Log likelihood	-331.3206	F-statistic	27.90230
Durbin-Watson stat	2.002667	Prob(F-statistic)	0.000000

Inverted AR Roots	.16 -.42i	.16+.42i

美国月利率数据的拟合－ARMA

- 从ACF图形来看，似乎是二阶截尾的，可以考虑MA(2)模型
- 输入 ls dr c ma(1) ma(2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.004486	0.029275	0.153252	0.8783
MA(1)	0.334211	0.047184	7.083125	0.0000
MA(2)	-0.127170	0.047184	-2.695172	0.0073

R-squared	0.112868	Mean dependent var	0.004569
Adjusted R-squared	0.108854	S.D. dependent var	0.542089
S.E. of regression	0.511735	Akaike info criterion	1.504697
Sum squared resid	115.7475	Schwarz criterion	1.532325
Log likelihood	-331.7951	F-statistic	28.11750
Durbin-Watson stat	1.994661	Prob(F-statistic)	0.000000

Inverted MA Roots	.23	-.56

ARMA模型p,q阶数的确定

- 可以采用试错法（Trial and Error）方法加以选取使得SBC值最小的p与q值
- Schwartz（1978）提出的SBC（Schwarz Bayesian Criterion）准则
- $SBC(M)=n\ln\sigma_{\varepsilon}^{2}+M\ln n$
- 其中σ_{ε}^{2}为最大似然估计量，M为模型中参数的个数，n为有效的观测指数
- 在试错过程中，p与q值阶数均限制在3（含3）以下
- 该方法较为客观

美国月利率数据的拟合－ARMA

不同p,q取值时，模型拟合的SBC值。当(p,q)=(3,3)时模型拟合的SBC值最小。

P Q	0	1	2	3
0	1.624680	1.563960	1.537070	1.553134
1	1.530077	1.534262	1.550752	1.556467
2	1.532325	1.541504	1.556771	1.563041
3	1.545586	1.554925	1.564360	1.485790

美国月利率数据的拟合－ARMA

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.042584	0.015807	2.694085	0.0073
AR(1)	-0.311711	0.101515	-3.070596	0.0023
AR(2)	-0.896754	0.024629	-36.41083	0.0000
AR(3)	-0.477184	0.100714	-4.737997	0.0000
MA(1)	0.644363	0.080750	7.979745	0.0000
MA(2)	0.955682	0.011524	82.93066	0.0000
MA(3)	0.753795	0.084310	8.940730	0.0000

R-squared	0.203930	Mean dependent var	0.004321
Adjusted R-squared	0.192949	S.D. dependent var	0.543834
S.E. of regression	0.488559	Akaike info criterion	1.420996
Sum squared resid	103.8299	Schwarz criterion	1.485790
Log likelihood	-307.0400	F-statistic	18.57234
Durbin-Watson stat	1.975533	Prob(F-statistic)	0.000000

ARIMA模型参数的限制

- 对于AR(1)、AR(2)模型，系数ϕ_1、ϕ_2的绝对值通常限制低于1，该限制称为稳定性界限，若超出该界限，则该时序不是自回归。可能存在单位根或趋势
- 对于MA(1)、MA(2)模型，系数θ_1、θ_2的绝对值限制低于1，该限制称为可逆性界限，若超出该界限，则该模型就不稳定
- 稳定性界限：

  AR(1)　$-1<\phi_1<1$，

  AR(2)　$-1<\phi_{1,2}<1$，$\phi_1+\phi_2<1$，$\phi_2-\phi_1<1$
- 可逆性界限

  MA(1)　$-1<\theta_1<1$

  MA(2)　$-1<\theta_{1,2}<1$，$\theta_1+\theta_2<1$，$\theta_2-\theta_1<1$

残差项的白噪音检验

- 当预测模式拟合后，对残差值可以算出各滞后期的自相关，将残差值的自相关系数经Q检验以决定残差项是否符合白噪音假设
- 当检验后如符合白噪音假设，则表示所选用的模型合适，可以此模型进行预测
- 若经模型后不符合白噪音假设，即有一型态存在，表示所选用的模型不合适，需要重新拟合预测模型，直至通过Q检验为止

残差项的白噪音检验

- 通常使用Ljung-Box（1978）提出的修正的Q检验，作为残差项白噪音检验的方法

$$\tilde{Q}(K) = n(n+2)\sum_{j=1}^{k}(n-j)^{-1}\hat{\rho}_e^2(j) \sim \chi^2_{k-p-q}$$

k通常可取24或k~ln(n)。

其虛假設為：

$H_0: \{a_t\}_{t=1}^{n}$ 為白噪音

$H_1: \{a_t\}_{t=1}^{n}$ 不為白噪音

其拒絕域為 $\tilde{Q}(K) \ge \chi^2_{k-p-q}$，屬右尾之單尾檢定。

美国月利率数据的拟合－ARMA

- ls dr c ar(1) ar(2) ar(3) ma(1) ma(2) ma(3)
- new series et=resid
- View/correlogram/OK

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.008	0.008	0.0295	0.864
		2	0.028	0.028	0.3720	0.830
		3	-0.033	-0.033	0.8467	0.838
		4	-0.062	-0.063	2.5805	0.630
		5	-0.030	-0.027	2.9802	0.703
		6	-0.150	-0.148	13.042	0.042
		7	-0.068	-0.072	15.159	0.034
		8	0.110	0.115	20.679	0.008
		9	0.059	0.052	22.274	0.008
		10	-0.001	-0.032	22.274	0.014
		11	0.113	0.106	28.106	0.003
		12	-0.039	-0.049	28.808	0.004
		13	-0.102	-0.128	33.540	0.001
		14	0.162	0.220	45.546	0.000
		15	-0.106	-0.070	50.692	0.000
		16	0.134	0.110	58.955	0.000

促使残差项为白噪音的方法

- 就上述模型来看，残差序列存在高阶（6阶以上）序列自相关
- 当残差项无法通过白噪音检验时，以往解决此问题的方法通常是将残差拟合成ARIMA 模型，再代入原序列的ARIMA(p,d,q)模型，进而得到修正后的模型，但有时常会造成(p,q)阶数过高，过度拟合
- 另一种方法是先采用Box-Cox变换，并通过编程求得使SSE最小的转换参数，对原数据进行转换，以促使原数据方差稳定化并服从渐近常态分配，之后又对离群值加以修正，使得模型能顺利地通过白噪音检验

预测准确度的评估指标—MAPE

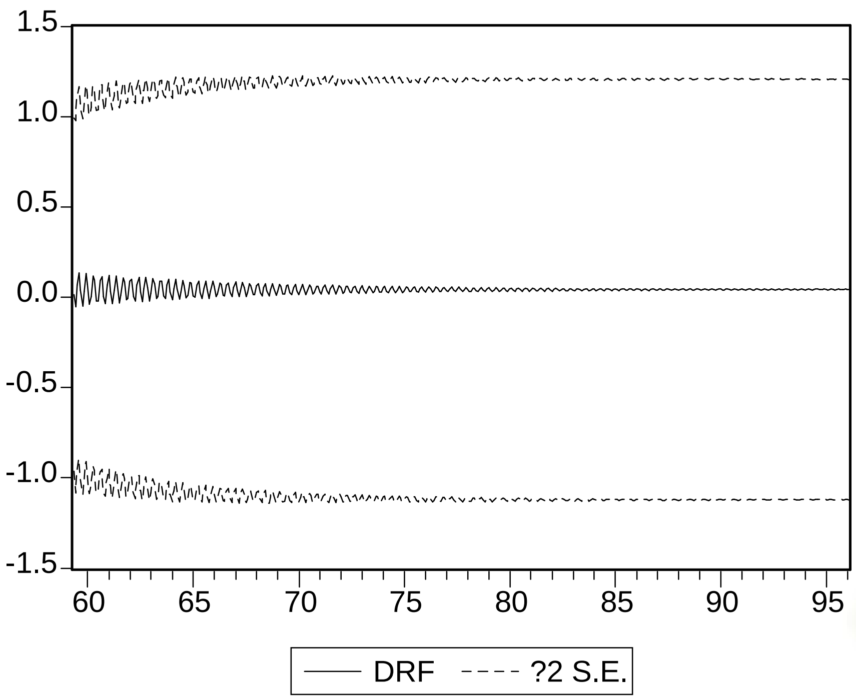
- 平均绝对值误差(Mean Absolute Percentage Error)，其公式为

$$\frac{\ell}{M}\sum_{t=1}^{M}\left|e_t\right|$$

- 其中，M为预测期数，l为预测第几期数
- 一般来说，MAPE因分母为实际值，所代表为百分比，不会因数值的大小而产生无可比性的问题
- MAPE评估预测准确度准则(Martin & Witt,1989)：MAPE <10 预测准确度极佳；10<MAPE<20 预测准确度好；20<MAPE<50 预测准确度普通；50<MAPE 预测准确度差

美国月利率数据的拟合－ARMA

- ls dr c ar(1) ar(2) ar(3) ma(1) ma(2) ma(3)
- Forecast/OK   ARMA(3,3)模型拟合情况较差

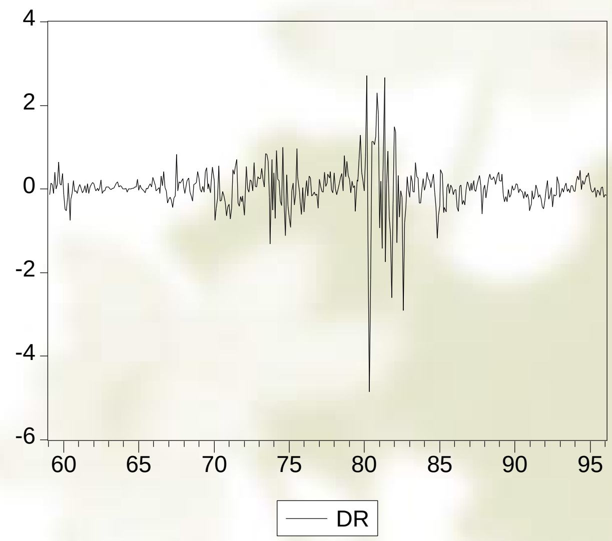


Forecast: DRF
Actual: DR
Forecast sample: 1959:01 1996:
Adjusted sample: 1959:05 1996:
Included observations: 442

Root Mean Squared Error	0.545672
Mean Absolute Error	0.307682
Mean Abs. Percent Error	175.9600
Theil Inequality Coefficient	0.921759
Bias Proportion	0.004881
Variance Proportion	0.905436
Covariance Proportion	0.089682

GARCH类模型

- ARIMA是对时序的均值进行预测的模型
- 而GARCH类模型是对时序的条件方差（波动率，volatility）进行预测的模型
- 其预测的出发点是波动率的群集现象（volatility clustering）以及非对称现象等



GARCH类模型的设定

- GARCH类模型包括两部分，一是条件均值的预测模型，一是条件方差的预测模型
- 标准GARCH（1,1）模型的设定为：

$$y_t = x_t\gamma + \epsilon_t$$
$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

- 均值方程包括外生变量的函数和一个误差项
- σ_t^2为方差的向前一期预测值，即条件方差。其函数包含三项：均值，上一期的波动率ε_{t-1}^2（ARCH项），上一期的预测方差σ_{t-1}^2（GARCH项）

GARCH类模型的设定

- GARCH(1,1)，前一个参数表示一阶GARCH项，后一个参数表示一阶ARCH项
- ARCH模型即为GARCH(0,1)模型，无GARCH项
- GARCH(p,q)模型设定为

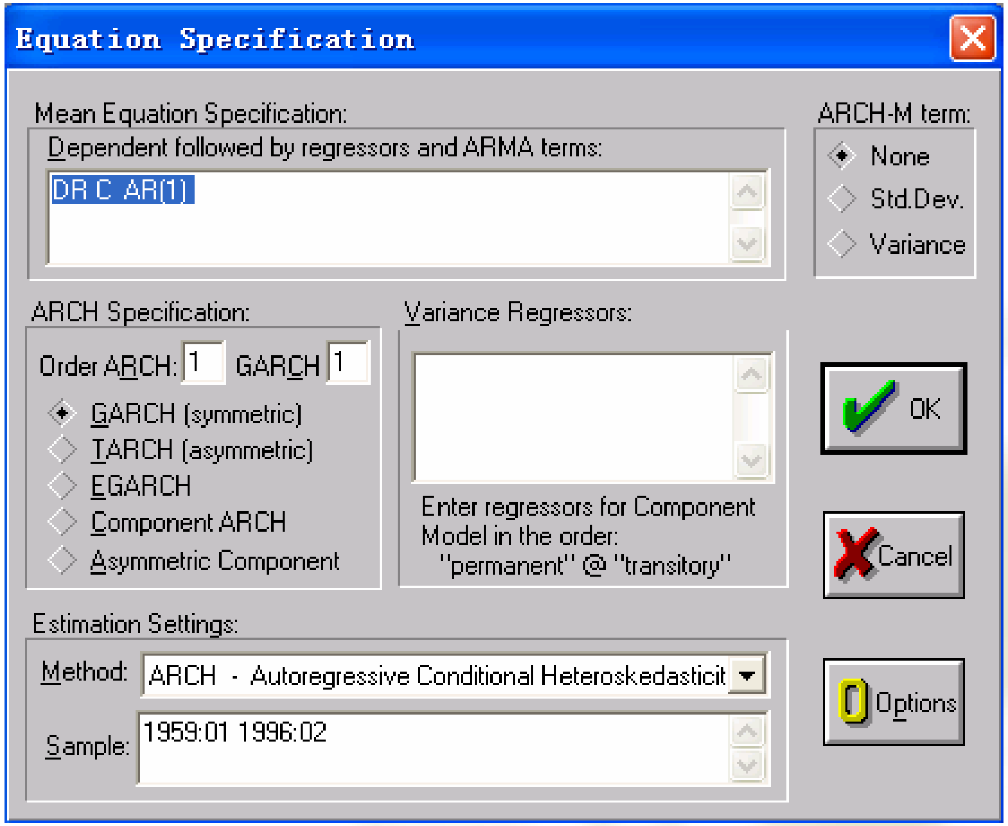
$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

- 还可以在方差模型中引入外生变量

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \pi z_t$$

美国月利率数据的拟合－GARCH

- Quick/Estimate Equation
- 在method中选择ARCH



美国月利率数据的拟合－GARCH

- 该模型设定说明均值方程为

  $dr_t = c + \phi_1 dr_{t-1} + \varepsilon_t$

- 方差方程为

  $\sigma_t^2 = \omega + \alpha\,\varepsilon_{t-1}^2 + \beta\,\sigma_{t-1}^2$

- 估计结果为

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.029142	0.013326	2.186921	0.0287
AR(1)	0.362682	0.043466	8.344131	0.0000
Variance Equation				
C	0.001140	0.000386	2.954294	0.0031
ARCH(1)	0.295118	0.040154	7.349718	0.0000
GARCH(1)	0.755002	0.027549	27.40614	0.0000

美国月利率数据的拟合

- View/Residual tests/Correlogram-Q-Statistics－说明残差不存在序列自相关

Date: 07/12/05 Time: 14:08
Sample: 1959:03 1996:02
Included observations: 444
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.037	0.037	0.6266	
		2	0.000	-0.001	0.6266	0.429
		3	0.031	0.031	1.0515	0.591
		4	0.006	0.004	1.0705	0.784
		5	0.086	0.086	4.4439	0.349
		6	0.017	0.010	4.5779	0.470
		7	-0.073	-0.074	6.9858	0.322
		8	0.037	0.037	7.6001	0.369
		9	0.045	0.041	8.5080	0.385
		10	0.011	0.004	8.5605	0.479
		11	0.070	0.067	10.817	0.372

美国月利率数据的拟合－GARCH

- View/Residual tests/ARCH-LM Test－残差不存在ARCH效应

ARCH Test:

F-statistic	0.000677	Probability	0.979254
Obs*R-squared	0.000680	Probability	0.979195

- 说明用AR(1)-GARCH(1,1)模型拟合dr较为充分

ARCH-M模型的设定

- ARCH-M模型将条件方差引入均值方程（ARCH-in-Mean, Engle, Lilien, Robins, 1987），模型设定为

$$y_t = x_t'\gamma + \sigma_t^2\bar{\gamma} + \epsilon_t$$

- 该设定的经济意义为，资产的期望收益与期望风险相关，期望风险系数的估计值可视为风险－收益的抵换系数（risk-return tradeoff）

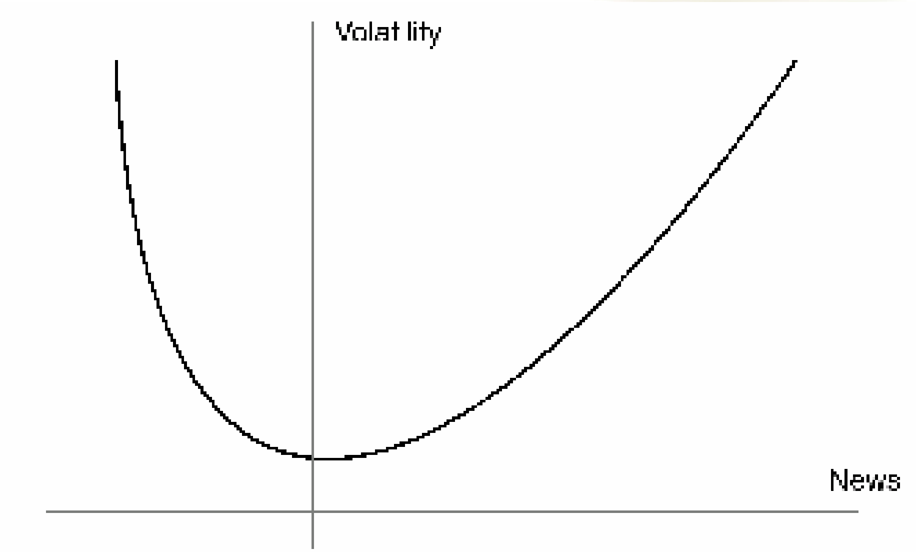
美国月利率数据的拟合－ARCH-M

- 在均值方程中加入条件方差或条件标准差，似乎均不显著

	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	-0.097304	0.118360	-0.822102	0.4110
C	0.025831	0.013163	1.962381	0.0497
AR(1)	0.357442	0.044114	8.102665	0.0000
Variance Equation				
C	0.001108	0.000380	2.911986	0.0036
ARCH(1)	0.284490	0.038495	7.390336	0.0000
GARCH(1)	0.759017	0.026753	28.37087	0.0000

非对称GARCH模型

- 通常在金融市场中可以发现，熊市的波动率高于牛市，Engle and Ng（1993）应用非对称新闻影响曲线来说明该现象
- EViews可以应用两类模型来估计对波动率的非对称冲击：TARCH和EGARCH模型



TGARCH模型的设定

- TGARCH模型或门限GARCH模型，由Zakoian(1990)和Glosten,Jaganathan and Runkle(1993)提出，模型设定为

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma\epsilon_{t-1}^2 d_{t-1} + \beta\sigma_{t-1}^2$$

- 其中，当ε $_t$<0时 **d_t=1**，否则 **d_t=0**
- 在该模型中，好消息和坏消息对条件方差有不同影响
- TGARCH(p,q)模型的设定为

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i\epsilon_{t-i}^2 + \gamma\epsilon_{t-1}^2 d_{t-1} + \sum_{j=1}^{p} \beta_j\sigma_{t-j}^2$$

美国月利率数据的拟合－TGARCH

- TGARCH模型估计结果显示存在显著的非对称效应

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.039645	0.014757	2.686565	0.0072
AR(1)	0.342005	0.044121	7.751554	0.0000
Variance Equation				
C	0.000911	0.000334	2.723253	0.0065
ARCH(1)	0.348668	0.056829	6.135335	0.0000
(RESID<0)*ARCH(1)	-0.202200	0.051978	-3.890113	0.0001
GARCH(1)	0.794953	0.027635	28.76667	0.0000

EGARCH模型的设定

- EGARCH或指数GARCH模型由Nelson(1991)提出，其条件方差设定为

$$\log(\sigma_t^2) = \omega + \beta \log(\sigma_{t-1}^2) + \alpha \left| \frac{\epsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}}$$

- EGARCH(p,q)模型设定为

$$\log(\sigma_t^2) = \omega + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2) + \sum_{i=1}^{q} \left(\alpha_i \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right)$$

美国月利率数据的拟合－EGARCH

- EGARCH模型估计结果显示存在显著的非对称效应

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.058374	0.012339	4.730849	0.0000
AR(1)	0.349260	0.045362	7.699361	0.0000
Variance Equation				
C	-0.350487	0.040163	-8.726627	0.0000
|RES|/SQR[GARCH](1	0.388859	0.041404	9.391873	0.0000
RES/SQR[GARCH](1)	0.108497	0.024151	4.492539	0.0000
EGARCH(1)	0.971998	0.007018	138.5056	0.0000

Component ARCH模型的设定

- Component ARCH模型将波动率分为永久性和暂时性两部分，其设定为

$$\sigma_t^2 - q_t = \alpha(\epsilon_{t-1}^2 - q_{t-1}) + \beta(\sigma_{t-1}^2 - q_{t-1})$$
$$q_t = \omega + \rho(q_{t-1} - \omega) + \phi(\epsilon_{t-1}^2 - \sigma_{t-1}^2).$$

- 通常假设0.99< ρ <1，Component ARCH模型近似于GARCH(2,2)
- 非对称Component 模型将其与TGARCH模型相结合，设定为

$$y_t = x_t\pi + \epsilon_t$$
$$q_t = \omega + \rho(q_{t-1} - \omega) + \phi(\epsilon_{t-1}^2 - \sigma_{t-1}^2) + \theta_1 z_{1t}$$
$$\sigma_t^2 - q_t = \alpha(\epsilon_{t-1}^2 - q_{t-1}) + \gamma(\epsilon_{t-1}^2 - q_{t-1})d_{t-1} + \beta(\sigma_{t-1}^2 - q_{t-1}) + \theta_2 z_{2t}$$

美国月利率数据的拟合－Component GARCH

- Component GARCH模型拟合结果说明暂时性部分不显著

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.020775	0.026090	0.796300	0.4259
AR(1)	0.212045	0.057725	3.673381	0.0002
Variance Equation				
Perm: C	0.168076	0.023313	7.209434	0.0000
Perm: [Q-C]	0.833926	0.023138	36.04062	0.0000
Perm: [ARCH-GARCH]	0.200350	0.064083	3.126408	0.0018
Tran: [ARCH-Q]	-0.032367	0.071472	-0.452867	0.6506
Tran: [GARCH-Q]	-0.285524	1.887064	-0.151306	0.8797