Conditional Volatility Models

Econ 147

UCLA

Version 1.1

Volatility Modeling

- In finance, one of the most important principle is "risk-return trade off".
  - return is typically what we expect from the investment, hence $E[r_t]$ (first moment)
    - or $E[r_t - r_f]$ where r_f is a risk-free rate (constant): risk premium
  - risk is typically measured by volatility, hence $Var(r_t)$ (second moment)
- **Reading:** Course slides and Supplemental Notes
- Optional: Chapter 3 (**Heteroscedastic Volatility Models**) in Fan and Yao's book
- Optional: Chapter 18 (**GARCH Models**) in Ruppert's book

Volatility Modeling

- In economics and finance, what we really care about is the "conditional moment"
  - $E\left[r_t|\mathcal{F}_{t-1}\right]$: conditional mean return, sometimes modelled by a regression framework

$$E\left[r_t|\mathcal{F}_{t-1}\right] = \alpha + \beta x_{t-1}$$

  - $Var\left(r_t|\mathcal{F}_{t-1}\right)$: conditional volatility, how to model?

Constant Expected Return (CER) Model

$$
\begin{gathered}
r_{it} = \mu_i + \epsilon_{it} \quad t = 1, \cdots, T; \quad i = 1, \cdots N \\
\epsilon_{it} \sim \text{iid } N(0, \sigma_i^2) \\
\text{cov}(\epsilon_{it},\ \epsilon_{jt}) = \sigma_{ij},\ \text{cor}(\epsilon_{it},\ \epsilon_{jt}) = \rho_{ij} \\
\text{cov}(\epsilon_{it},\ \epsilon_{js}) = 0 \quad t \neq s,\ \text{for all } i,\ j
\end{gathered}
$$

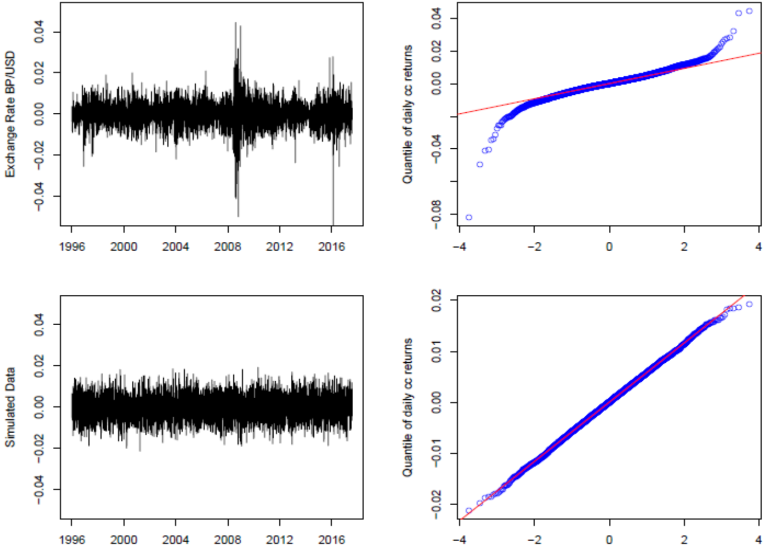
- CER model, a.k.a. iid Normal model
- We studied probability and statistical inference for this model, such as
  - Point Estimation
  - Interval Estimation (Confidence Interval)
  - Hypothesis Testing

Parameters of CER Model

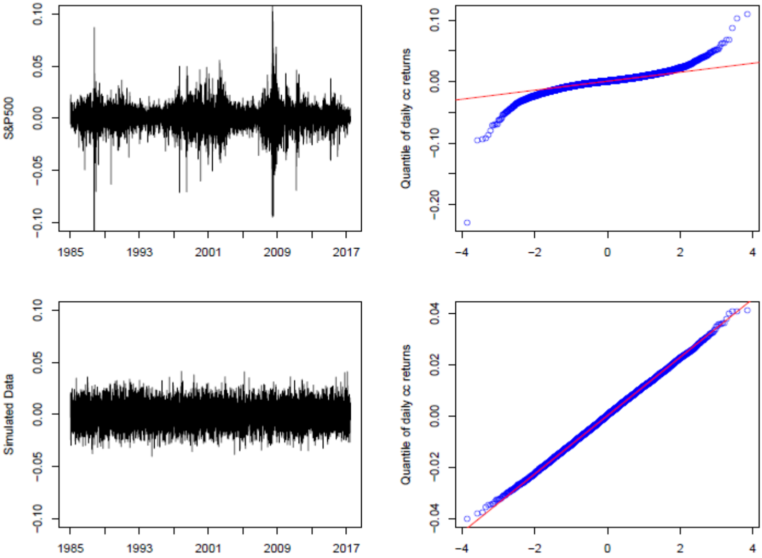
$$
\begin{aligned}
\mu_i &= E[r_{it}] \\
\sigma_i^2 &= \mathrm{var}(r_{it}) \\
\sigma_{ij} &= \mathrm{cov}(r_{it}, r_{jt}) \\
\rho_{ij} &= \mathrm{cor}(r_{it}, r_{jt})
\end{aligned}
$$

- Second moments of this model is time-invariant (constant), i.e., does not change with the time t
- Since $\epsilon_{it} \sim \mathrm{iid}\ N(0, \sigma_i^2)$, $\mathrm{var}(r_{it}|\mathcal{F}_{t-1}) = E(\epsilon_{it}^2|\mathcal{F}_{t-1}) = E(\epsilon_{it}^2) = \sigma_i^2$
- Implications from the CER Model:
  - normal distribution
  - constant variance over time
- are these implications realistic? NO!

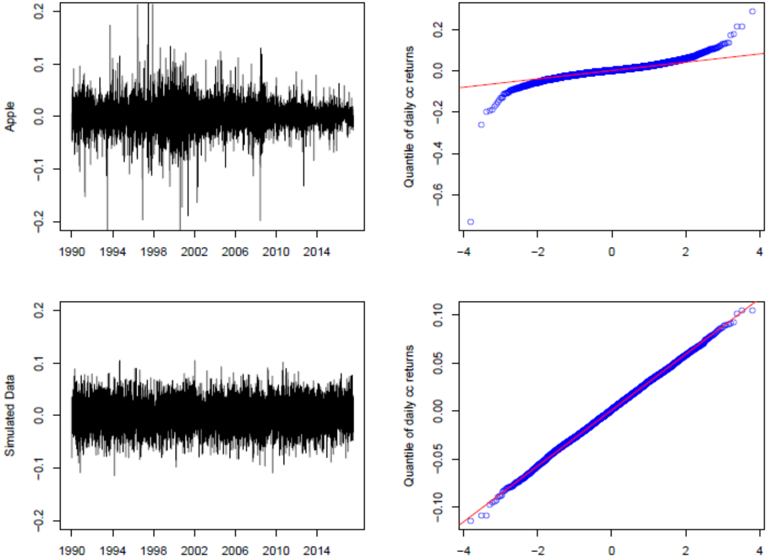
Daily CC Returns of Exchange Rate BP/USD (1996/1/2 to 2017/12/22)



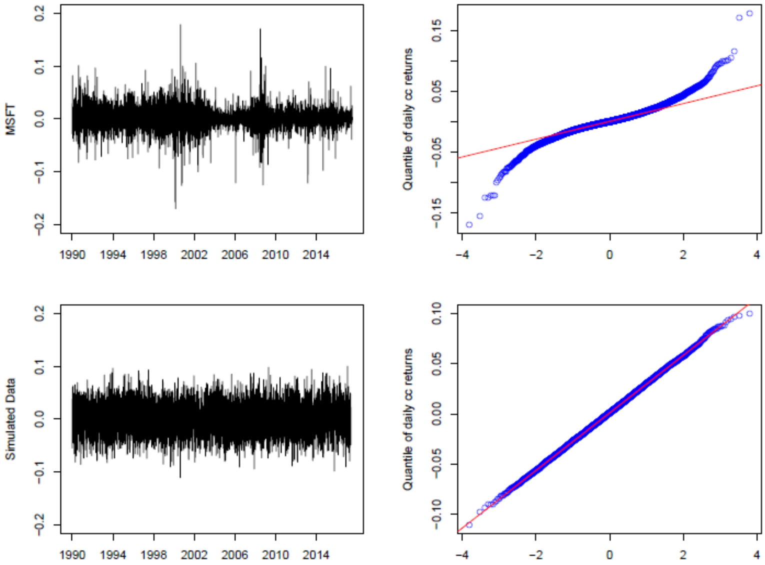
Daily CC Returns of S&P500 (1985/1/2 to 2017/12/22)



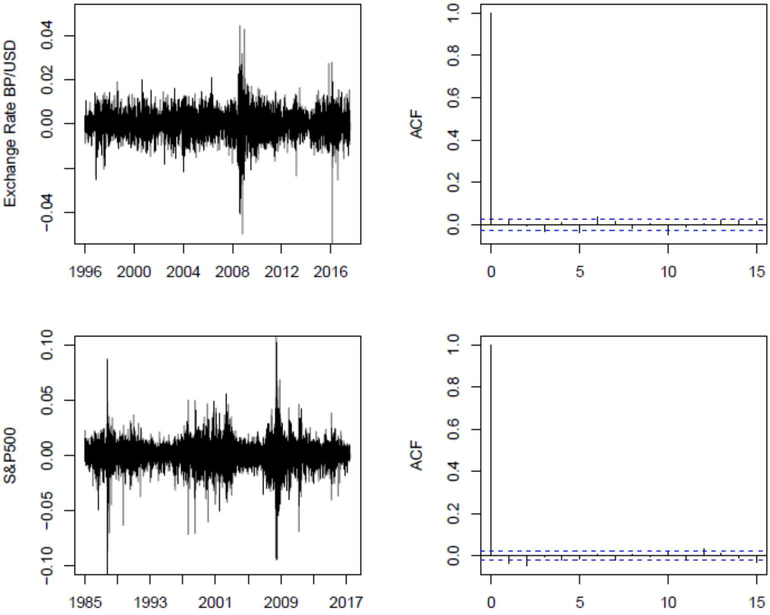
Daily CC Returns of Apple Shares (1990/1/2 to 2017/12/29)



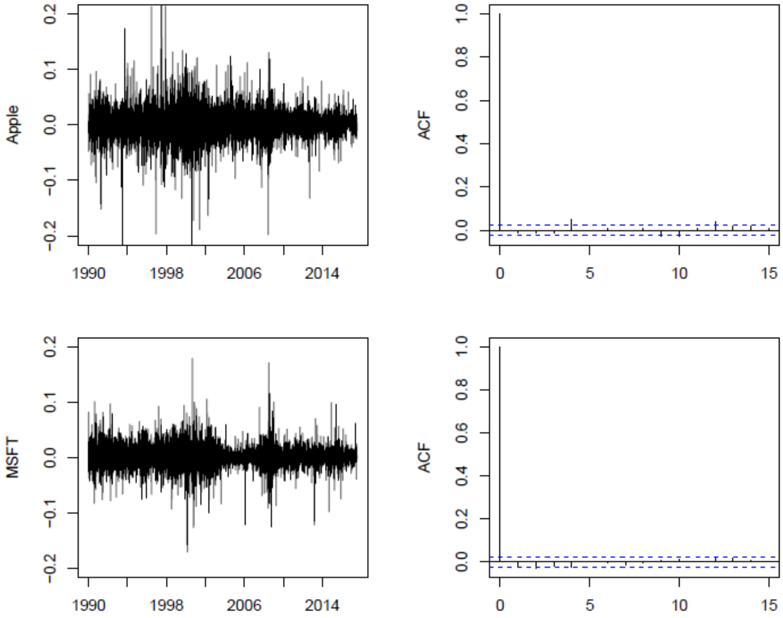
Daily CC Returns of Microsoft Shares (1990/1/2 to 2017/12/29)



Daily CC Returns of Exchange Rate and S&P500



Daily CC Returns of Apple and Microsoft Shares



Jarque Bera Test (H_0: $r_{it} \sim iid\ N(\mu_i, \sigma_i^2)$)				
	BP/USD	S&P500	Apple	Microsoft
JB Statistic	30953	285800	1222400	10578
Critical Value	5.9915	5.9915	5.9915	5.9915
Test Result	reject H_0	reject H_0	reject H_0	reject H_0

- H_0 of normal distribution is clearly rejected.
    - the main source of this rejection is, again, from outliers.

- Daily (cc) returns on Exchange Rate (BP/USD), S&P 500, Apple and Microsoft:
  - often exhibit volatility clustering; periods of high and low volatility
  - not clear evidence of autocorrelation: but not independent (WHY?)
  - more outliers than normal (Gaussian distributions): heavy tailedness
- IID Normal (Constant Expected Return) model does not capture the empirical property of these financial data
- proper modeling of time-varying volatility is of importance
- We introduce some popular class of volatility modeling that allows $\sigma_t^2 = \mathrm{var}(r_t|\mathcal{F}_{t-1})$ is time varying in the conditional sense (see below)
  - this is called the *conditional volatility modeling*

Stylized Facts on stock return

- Nearly mds: $E\left[r_t|\mathcal{F}_{t-1}\right] \simeq 0$
- Volatility clustering; periods of high and low volatility
  - persistent volatility modeling
- No (or weak) autocorrelation
- Heavier tail than normal distribution

ARCH(1) model

- Let $\mu_t = E\left[r_t | \mathcal{F}_{t-1}\right]$. Then
$$r_t = \mu_t + (r_t - \mu_t) = \mu_t + X_t \text{ where } X_t = r_t - \mu_t$$
- ARCH(1):
$$\begin{aligned} X_t &= \sigma_t e_t,\ e_t \sim iid\ N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 X_{t-1}^2 \end{aligned}$$
- $\omega > 0$, $\alpha_1 \geq 0$ are imposed usually
- Since $\mu_t = E\left[r_t | \mathcal{F}_{t-1}\right] \simeq 0$, $r_t \simeq X_t$
- More often we write
$$\begin{aligned} r_t &= \sigma_t e_t,\ e_t \sim iid\ N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 \end{aligned}$$
- In the rest of the slides, we maintain the assumption that $\mu_t = 0$ (or we can write $r_t = r_t - \mu_t$).

ARCH(1) model

Motivated by earlier discussion:

$$\begin{aligned} r_t &= \sigma_t e_t, \; e_t \sim iid \; N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 \end{aligned}$$

or equivalently

$$r_t = \sqrt{\omega + \alpha_1 r_{t-1}^2} e_t, \; e_t \sim iid \; N(0,1)$$

- $\omega > 0$, $\alpha_1 \geq 0$ are imposed usually
- σ_t^2 is known at time $t-1$ (hence included in $\mathcal{F}_{t-1}$)

ARCH(1) model

- The volatility at time t (σ_t^2) depends on the past squared return (r_{t-1}^2) in the *autoregressive* way.
- **A**uto**R**egressive **C**onditional **H**eteroskedasticity - ARCH (Robert Engle, 1982)
- Huge success in finance/econometrics/statistics.
  - Engle received Nobel Prize from this contribution

ARCH(1) model properties: mds

- From

$$
\begin{aligned}
r_t &= \sigma_t e_t, \;\; e_t \sim iid \;\; N(0,1) \\
\sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2
\end{aligned}
$$

  - $E\left[r_t | \mathcal{F}_{t-1}\right] = 0$ (mds), hence $cov\left(r_t, r_s\right) = 0$ for all $t \neq s$.

ARCH(1) model properties: persistent volatility

- From

$$\begin{aligned} r_t &= \sigma_t e_t, \ e_t \sim iid\ N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 \end{aligned}$$

- Let $v_t = r_t^2 - \sigma_t^2 = r_t^2 - E\left[r_t^2 | \mathcal{F}_{t-1}\right]$, then $E\left[v_t | \mathcal{F}_{t-1}\right] = 0$ (mds)
  - $v_t = r_t^2 - E\left[r_t^2 | \mathcal{F}_{t-1}\right]$ is, by construction, unexpected part from $\mathcal{F}_{t-1}$, so naturally time series error/innovation/news
- By adding v_t to both sides,

$$\sigma_t^2 + v_t = \omega + \alpha_1 r_{t-1}^2 + v_t.$$

Since $v_t = r_t^2 - \sigma_t^2$, we have

$$r_t^2 = \omega + \alpha_1 r_{t-1}^2 + v_t$$

ARCH(1) model properties: persistent volatility

- AR(1) representation in "squares"

$$r_t^2 = \omega + \alpha_1 r_{t-1}^2 + v_t.$$

- If we assume $\alpha_1 < 1$, then r_t^2 is stationary AR(1) process.
  - $E\left[r_t^2\right] = \frac{\omega}{1-\alpha_1}$
  - $Corr(r_t^2, r_{t-j}^2) = \alpha_1^{|j|}$
- Squared return (volatility proxy) can be persistent: $\alpha_1 \simeq 1$ (see below)

ARCH(1) model properties: heavy tailedness

- ARCH(1) returns $r_t = \sigma_t e_t$, $e_t \sim iid\ N(0,1)$ has heavier tails than normal distribution.
- Intuitively, outlier occurs when the variance is large, so it gets even bigger.

ARCH(1) model properties: heavy tailedness

- To prove, we need to show

$$kurt(r_t) = \frac{E\left[r_t^4\right]}{\left(E\left[r_t^2\right]\right)^2} \geq 3 = kurt(e_t) = \frac{E\left[e_t^4\right]}{\left(E\left[e_t^2\right]\right)^2} = E\left[e_t^4\right]$$

- Note that

$$\begin{aligned} E\left[r_t^4 | \mathcal{F}_{t-1}\right] &= E\left[\sigma_t^4 e_t^4 | \mathcal{F}_{t-1}\right] \\ &= \sigma_t^4 E\left[e_t^4\right] \\ &= \sigma_t^4 \times 3 \\ &= \left(\sigma_t^2\right)^2 \times 3 \\ &= \left(E\left[r_t^2 | \mathcal{F}_{t-1}\right]\right)^2 \times 3 \end{aligned}$$

ARCH(1) model properties: heavy tailedness

- Using Jensen's Inequality: $E\left[Y^2\right] \geq \left(E\left[Y\right]\right)^2$,

$$E\left[\left(E\left[r_t^2|\mathcal{F}_{t-1}\right]\right)^2\right] \geq \left(E\left[E\left[r_t^2|\mathcal{F}_{t-1}\right]\right]\right)^2 = \left(E\left[r_t^2\right]\right)^2$$

thus $E\left[\left(E\left[r_t^2|\mathcal{F}_{t-1}\right]\right)^2\right] \times 3 \geq \left(E\left[r_t^2\right]\right)^2 \times 3$.

- Now

$$E\left[E\left[r_t^4|\mathcal{F}_{t-1}\right]\right] = E\left[\left(E\left[r_t^2|\mathcal{F}_{t-1}\right]\right)^2\right] \times 3 \geq \left(E\left[r_t^2\right]\right)^2 \times 3$$

therefore

$$\frac{E\left[r_t^4\right]}{\left(E\left[r_t^2\right]\right)^2} \geq 3.$$

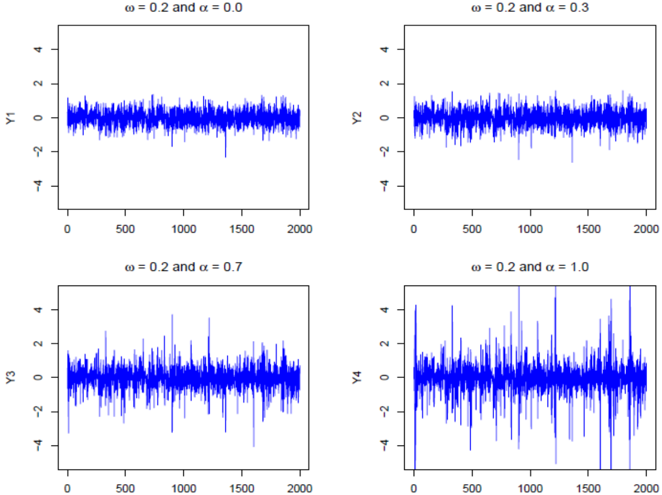
ARCH(p) model

- We could introduce ARCH(p) model:

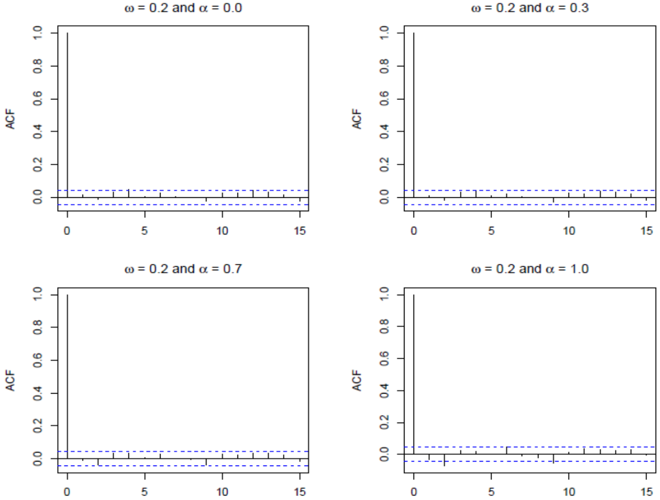
$$
\begin{aligned}
r_t &= \sigma_t e_t, \; e_t \sim iid \; N(0,1) \\
\sigma_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2
\end{aligned}
$$

- Alternatively, to have parsimonious model, we introduce GARCH(1,1)
  - remind the relation of AR(p) and ARMA(1,1)

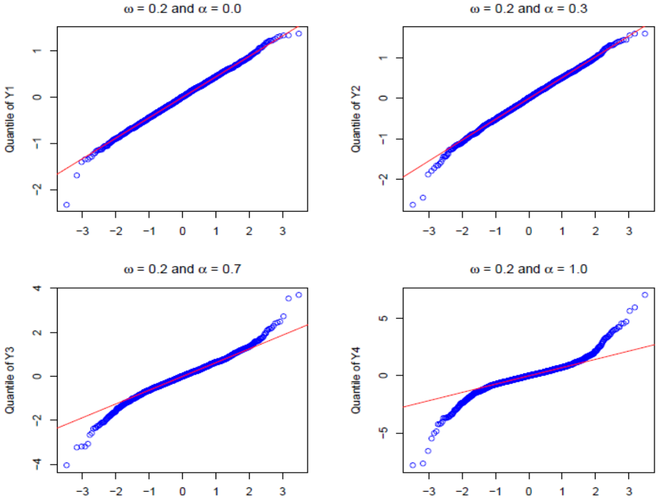
$$\text{ARCH(1)}: \; r_t = \sqrt{\omega + \alpha r_{t-1}^2} e_t, \; e_t \sim iid \; N(0,1)$$



$$\text{ARCH(1)}: \ r_t = \sqrt{\omega + \alpha r_{t-1}^2}\, e_t, \ e_t \sim iid \ N(0,1)$$



$$\text{ARCH(1)}: \; r_t = \sqrt{\omega + \alpha r_{t-1}^2}\, e_t, \; e_t \sim iid \;\; N(0,1)$$



Jarque Bera Test for ARCH(1) (H_0: $r_{it} \sim iid\ N(\mu_i, \sigma_i^2)$)				
	$Y1$	$Y2$	$Y3$	$Y4$
	$(\alpha = 0)$	$(\alpha = .3)$	$(\alpha = .7)$	$(\alpha = 1)$
Statistic	2.1857	180.497	579265	2958117
C.V.	5.9915	5.9915	5.9915	5.9915
Rej. Prob.	0.07	0.94	1	1

Note: the above results are based on 100 replications

GARCH(1,1) model

- As an alternative ARCH(p):

$$\begin{aligned} r_t &= \sigma_t e_t, \; e_t \sim iid \; N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned}$$

  - $\omega > 0$, $\alpha_1 \geq 0$ and $\beta_1 \geq 0$ are imposed usually
- **G**eneralized **A**uto**R**egressive **C**onditional **H**eteroskedasticity - GARCH
- Even more successful than ARCH
- It has been an "industry standard" to use GARCH(1,1) for stock return data
  - Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. *Journal of applied econometrics*, 20(7), 873-889.

GARCH(1,1) model

- By similar procedure, with $r_t = \sigma_t e_t$, $e_t \sim iid(0,1)$,

$$\begin{aligned}
\sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\
r_t^2 - \sigma_t^2 + \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + r_t^2 - \sigma_t^2 \\
&= \omega + (\alpha_1 + \beta_1)\, r_{t-1}^2 \\
&\quad -\beta_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + r_t^2 - \sigma_t^2
\end{aligned}$$

so

$$r_t^2 = \omega + (\alpha_1 + \beta_1)\, r_{t-1}^2 + v_t - \beta_1 v_{t-1}$$

which can be further written as

$$\begin{aligned}
& r_t^2 - \frac{\omega}{1-\alpha_1-\beta_1} \\
= \;& (\alpha_1+\beta_1)\left(r_{t-1}^2 - \frac{\omega}{1-\alpha_1-\beta_1}\right) + v_t - \beta_1 v_{t-1}
\end{aligned}$$

- r_t^2 is ARMA (1,1) process.

- ARMA (1,1) model

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_t \sim MDS\ (0, \sigma_\varepsilon^2),\ \ |\phi| < 1$$

- Properties

$$E[Y_t] = \mu, \quad \operatorname{var}(Y_t) = \sigma_Y^2 = \frac{1+\theta^2+2\phi\theta}{1-\phi^2}\sigma_\varepsilon^2$$

$$\gamma_j = \operatorname{cov}(Y_t, Y_{t-j}) = \begin{cases} \frac{(1+\phi\theta)(\phi+\theta)}{1-\phi^2}\sigma_\varepsilon^2, & |j| = 1 \\ \phi\gamma_{j-1}, & j \geq 2 \\ \gamma_{-j}, & j \leq -2 \end{cases}$$

$$\rho_j = \operatorname{corr}(Y_t, Y_{t-j}) = \begin{cases} \frac{(1+\phi\theta)(\phi+\theta)}{1+\theta^2+2\phi\theta}, & |j| = 1 \\ \phi\rho_{j-1}, & j \geq 2 \\ \rho_{-j}, & j \leq -2 \end{cases}$$

- Set $\mu = \omega(1-\alpha_1-\beta_1)^{-1}$, $\phi = \alpha_1 + \beta_1$ and $\theta = -\beta_1$ to get the mean, variance and covariance of $\{r_t^2\}_t$.

GARCH(1,1) model: volatility dynamics

- When $(\alpha_1 + \beta_1) < 1$, we can show

$$E[r_t^2] = \frac{\omega}{1 - \alpha_1 - \beta_1},$$

$$\mathrm{var}(r_t^2) = E\left[v_t^2\right]\left\{1 + \frac{\alpha_1^2}{1 - (\alpha_1 + \beta_1)^2}\right\},$$

$$\gamma_j = \mathrm{cov}(r_t^2, r_{t-j}^2) = \begin{cases} \alpha_1 \frac{1-(\alpha_1+\beta_1)\beta_1}{1-(\alpha_1+\beta_1)^2} E\left[v_t^2\right], & |j| = 1 \\ (\alpha_1 + \beta_1)\gamma_{j-1}, & |j| \geq 2 \end{cases},$$

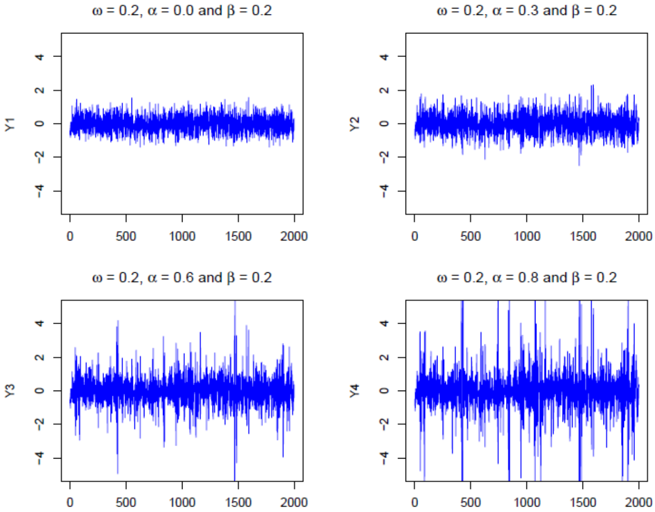
$$\rho_j = (\alpha_1 + \beta_1)\rho_{j-1} \text{ for } |j| \geq 2,$$

where ρ_j here is ACF of r_t^2 NOT r_t

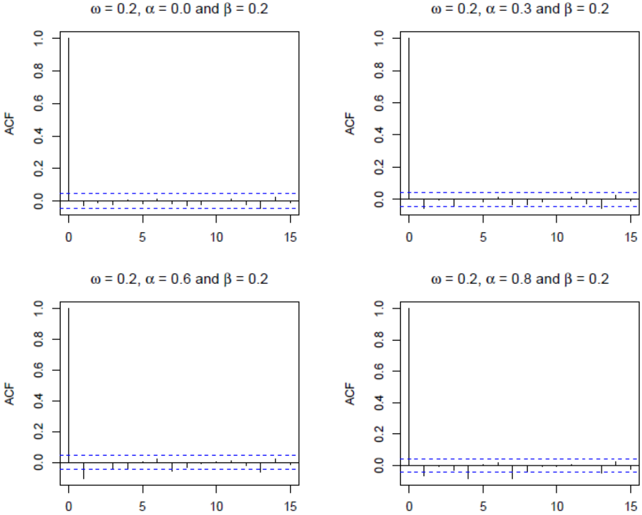
GARCH(1,1) model

- Many stylized facts of financial returns can be properly captured:
  - near mds so zero (or weak) autocorrelation
  - persistent volatility dynamics (clustering): $\alpha_1 + \beta_1 \simeq 1$ (see below)
  - heavy tailedness
- Only need a few parameters - parsimonious modeling

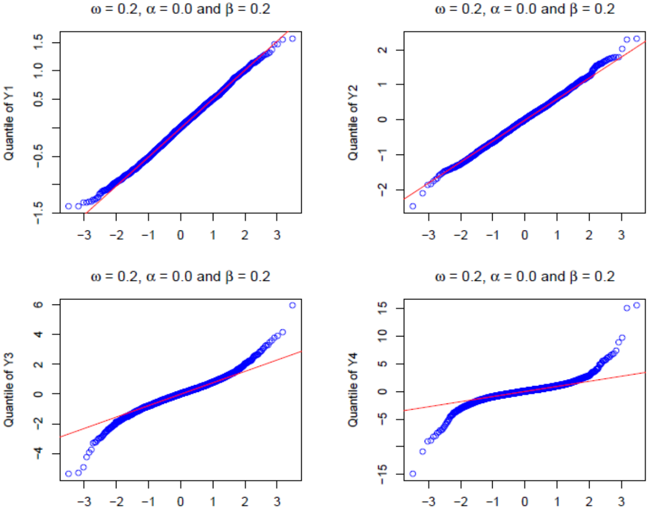
$$\text{GARCH(1,1)}: \; r_t = \left(\omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2\right)^{1/2} e_t, \; e_t \sim \; iid \; N(0,1)$$



GARCH(1,1): $r_t = \left(\omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2\right)^{1/2} e_t$, $e_t \sim iid\ N(0,1)$



$$\text{GARCH(1,1):}\ \ r_t = \left(\omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2\right)^{1/2} e_t,\ e_t \sim iid\ \ N(0,1)$$



Jarque Bera Test for GARCH(1,1) (H_0: $r_{it} \sim iid\ N(\mu_i, \sigma_i^2)$)				
	$Y1$	$Y2$	$Y3$	$Y4$
	$(\alpha = 0)$	$(\alpha = .3)$	$(\alpha = .6)$	$(\alpha = .8)$
Statistic	1.8002	72.242	19839.7	418870
C.V.	5.9915	5.9915	5.9915	5.9915
Rej. Prob.	0.05	0.9	1	1

Note: the above results are based on 100 replications

Confidence Interval

- Consider GARCH(1,1)

$$\begin{aligned} r_t &= \sigma_t e_t, \;\; e_t \sim iid\;\; N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned}$$

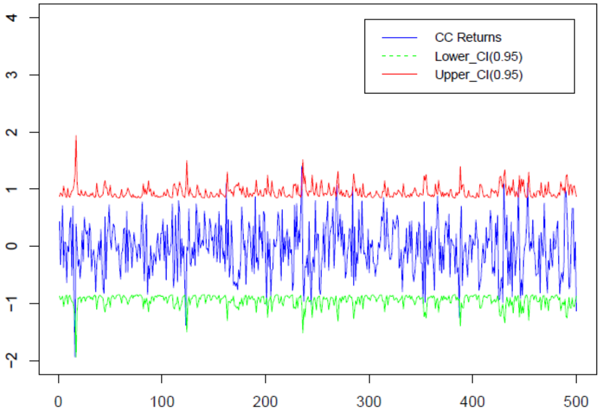
- By definition, r_t/σ_t is a standard normal random variable
- Therefore

$$\Pr\left(-q_{1-\alpha/2}^Z \leq r_t/\sigma_t \leq q_{1-\alpha/2}^Z\right) = 1-\alpha$$

where $q_{1-\alpha/2}^Z$ denotes the $100\,(1-\alpha/2)$ quantile of standard normal

- The $(1-\alpha)$-confidence interval of r_t is: $\left[-\sigma_t q_{1-\alpha/2}^Z,\;\; \sigma_t q_{1-\alpha/2}^Z\right]$
- For example, the 0.95-confidence interval of r_t is (approximately): $\left[-1.96\sigma_t,\;\; 1.96\sigma_t\right]$

$$r_t = \sqrt{0.1 + 0.2r_{t-1}^2 + 0.2\sigma_{t-1}^2}\, e_t,\; e_t \sim iid\ N(0,1)$$



Note: The empirical coverage probability is 0.97

Value at Risk (VaR)

- Suppose we have initial wealth W_0 invested in a risky asset
- With simple return R, the profit of this investment is $L_1 = W_0 \cdot R$ and its VaR_α is

$$q_\alpha^{L_1} = W_0 \cdot q_\alpha^R$$

  where q_α^R denotes the α quantile of R.
- With CC return r, the profit of this investment is $L_1 = W_0 \cdot (e^r - 1)$ and its VaR_α is

$$q_\alpha^{L_1} = W_0 \cdot (e^{q_\alpha^r} - 1)$$

  where q_α^r denotes the α quantile of r.

Value at Risk (VaR)

Lemma 2. Suppose that X is a random variable and $h(x)$ is a strictly increasing function of x. Define a new random variable $Y = h(X)$. Then

$$q_\alpha^Y = h(q_\alpha^X)$$

where q_α^Y and q_α^X denote the α quantiles of X and Y respectively.

- In GARCH models

$$r_t = \sigma_t e_t, \;\; e_t \sim iid \;\, N(0,1)$$

- The conditional distribution of r_t given σ_t is $N(0,1)$
- Given $\sigma_t > 0$, $r_t = \sigma_t e_t$ is a strictly increasing function of e_t.
- Therefore,

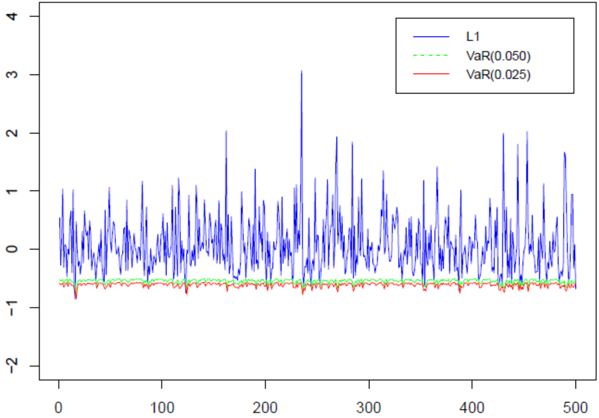
$$q_\alpha^{r_t} = \sigma_t q_\alpha^{e_t}$$

  where $q_\alpha^{r_t}$ and $q_\alpha^{e_t}$ denote the α quantiles of r_t and e_t (given $\mathcal{F}_{t-1}$) respectively.
- Therefore,

$$q_\alpha^{L_1} = W_0 \cdot (e^{q_\alpha^{r_t}} - 1) = W_0 \cdot (e^{\sigma_t q_\alpha^{e_t}} - 1)$$

- $q_\alpha^{e_t} = q_\alpha^{Z}$ is known since it is the quantile of the standard normal

$$L_{1,t} = \$1 * \left(e^{r_t} - 1\right), \text{ where } r_t = \sqrt{0.2 + 0.3 r_{t-1}^2}\, e_t,\ e_t \sim iid\ N(0,1)$$



Note: $W_0 = 1$. The relative frequencies of exceedances of $L_{1,t}$ over VaR(0.05) and VaR(0.025) are 0.050 and 0.016 respectively.

GARCH models

- The GARCH models are useful for
  - building confidence intervals: $\left[-\sigma_t q^Z_{1-\alpha/2},\ \sigma_t q^Z_{1-\alpha/2}\right]$
  - calculating the VaRs: $W_0 \cdot \left(e^{\sigma_t q^Z_\alpha} - 1\right)$
- The usefulness depends on the knowledge of σ_t which is unknown in reality
- Since $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, we can estimate σ_t^2 by estimating $(\omega, \alpha_1, \beta_1)$
- Estimation method - conditional maximum likelihood
  - facts: standard asymptotic results hold (consistency and asymptotic normality), hence usual t-test (or p-value based test) is available
  - most standard statistical package can accommodate GARCH model estimation

Estimation - conditional maximum likelihood

- From the ARCH(1) model:

$$\begin{aligned} r_t &= \sigma_t e_t, \;\; e_t \sim iid \;\; N(0,1) \\ \sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 \end{aligned}$$

the conditional distribution of $r_t|\mathcal{F}_{t-1} \sim N(0, \sigma_t^2)$ is given by

$$\begin{aligned} f\left(r_t|\mathcal{F}_{t-1};\theta\right) &= \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[-\frac{r_t^2}{2\sigma_t^2}\right] \\ &= \frac{1}{\sqrt{2\pi\left(\omega+\alpha_1 r_{t-1}^2\right)}} \exp\left[-\frac{r_t^2}{2\left(\omega+\alpha_1 r_{t-1}^2\right)}\right] \end{aligned}$$

where $\theta = (\omega, \, \alpha_1)$ denotes the parameter to be estimated.

Estimation - conditional maximum likelihood

- The conditional log-likelihood function for $\theta = (\omega, \alpha_1)$:

$$
\begin{aligned}
& \frac{1}{T}\log L_T(\theta) \\
= \quad & \frac{1}{T}\sum_{t=1}^{T}\log f\left(r_t|\mathcal{F}_{t-1};\theta\right) \\
= \quad & -\frac{1}{2}\log(2\pi) - \frac{1}{2T}\sum_{t=1}^{T}\log(\sigma_t^2) - \frac{1}{2T}\sum_{t=1}^{T}\frac{r_t^2}{\sigma_t^2} \\
= \quad & -\frac{1}{2}\log(2\pi) - \frac{1}{2T}\sum_{t=1}^{T}\log(\omega+\alpha_1 r_{t-1}^2) - \frac{1}{2T}\sum_{t=1}^{T}\frac{r_t^2}{\omega+\alpha_1 r_{t-1}^2}
\end{aligned}
$$

Estimation - conditional maximum likelihood

- FOC w.r.t ω:

$$
\begin{aligned}
\frac{\partial \log L_T(\theta)}{\partial \omega} &= \frac{\partial \log L_T(\theta)}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \omega} \\
&= \frac{1}{T} \sum_{t=1}^{T} \left(-\frac{1}{2\sigma_t^2} + \frac{r_t^2}{2\sigma_t^4} \right) \cdot 1 \\
&= \frac{1}{T} \sum_{t=1}^{T} \frac{1}{2\sigma_t^2} \left(\frac{r_t^2}{\sigma_t^2} - 1 \right)
\end{aligned}
$$

Estimation - conditional maximum likelihood

- FOC w.r.t α_1:

$$\begin{aligned}
\frac{\partial \log L_T(\theta)}{\partial \alpha_1} &= \frac{\partial \log L_T(\theta)}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \alpha_1} \\
&= \frac{1}{T} \sum_{t=1}^{T} \left(-\frac{1}{2\sigma_t^2} + \frac{r_t^2}{2\sigma_t^4} \right) \cdot r_{t-1}^2 \\
&= \frac{1}{T} \sum_{t=1}^{T} \frac{1}{2\sigma_t^2} \left(\frac{r_t^2}{\sigma_t^2} - 1 \right) r_{t-1}^2
\end{aligned}$$

The maximum likelihood estimator $\widehat{\theta}_T = (\widehat{\omega}_T, \widehat{\alpha}_{1,T})$ of $\theta = (\omega, \alpha_1)$ can be solved from the equations

$$
\begin{aligned}
\frac{\partial \log L_T(\widehat{\theta}_T)}{\partial \omega} &= 0 \text{ and} \\
\frac{\partial \log L_T(\widehat{\theta}_T)}{\partial \alpha_1} &= 0.
\end{aligned}
$$

Estimation - conditional maximum likelihood

- From the GARCH(1) model:

$$
\begin{aligned}
r_t &= \sigma_t e_t, \;\; e_t \sim iid \;\; N(0,1) \\
\sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\end{aligned}
$$

the conditional distribution of $r_t | \mathcal{F}_{t-1} \sim N(0, \sigma_t^2)$ is given by

$$
\begin{aligned}
f\left(r_t | \mathcal{F}_{t-1}; \theta\right) &= \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[-\frac{r_t^2}{2\sigma_t^2}\right] \\
&= \frac{\exp\left[-\frac{r_t^2}{2\left(\omega+\alpha_1 r_{t-1}^2+\beta_1\sigma_{t-1}^2\right)}\right]}{\sqrt{2\pi\left(\omega+\alpha_1 r_{t-1}^2+\beta_1\sigma_{t-1}^2\right)}}
\end{aligned}
$$

where $\theta = (\omega,\, \alpha_1,\, \beta_1)$ denotes the parameter to be estimated.

Estimation - conditional maximum likelihood

- The conditional log-likelihood function for $\theta = (\omega, \alpha_1, \beta_1)$:

$$
\begin{aligned}
& \frac{1}{T} \log L_T(\theta) \\
= \quad & \frac{1}{T} \sum_{t=1}^{T} \log f\left(r_t \middle| \mathcal{F}_{t-1}; \theta\right) \\
= \quad & -\frac{1}{2} \log(2\pi) - \frac{1}{2T} \sum_{t=1}^{T} \log(\sigma_t^2) - \frac{1}{2T} \sum_{t=1}^{T} \frac{r_t^2}{\sigma_t^2} \\
= \quad & -\frac{1}{2} \log(2\pi) - \frac{1}{2T} \sum_{t=1}^{T} \log(\omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2) \\
& -\frac{1}{2T} \sum_{t=1}^{T} \frac{r_t^2}{\omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2}
\end{aligned}
$$

Estimation - conditional maximum likelihood

- FOC w.r.t ω:

$$\begin{aligned}
\frac{\partial \log L_T(\theta)}{\partial \omega} &= \frac{\partial \log L_T(\theta)}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \omega} \\
&= \frac{1}{T} \sum_{t=1}^{T} \left(-\frac{1}{2\sigma_t^2} + \frac{r_t^2}{2\sigma_t^4} \right) \cdot 1 \\
&= \frac{1}{T} \sum_{t=1}^{T} \frac{1}{2\sigma_t^2} \left(\frac{r_t^2}{\sigma_t^2} - 1 \right)
\end{aligned}$$

Estimation - conditional maximum likelihood

- FOC w.r.t α_1:

$$\begin{aligned}
\frac{\partial \log L_T(\theta)}{\partial \alpha_1} &= \frac{\partial \log L_T(\theta)}{\partial \sigma_t^2} \frac{\partial \sigma_t^2}{\partial \alpha_1} \\
&= \frac{1}{T} \sum_{t=1}^{T} \left(-\frac{1}{2\sigma_t^2} + \frac{r_t^2}{2\sigma_t^4} \right) \cdot r_{t-1}^2 \\
&= \frac{1}{T} \sum_{t=1}^{T} \frac{1}{2\sigma_t^2} \left(\frac{r_t^2}{\sigma_t^2} - 1 \right) r_{t-1}^2
\end{aligned}$$

Estimation - conditional maximum likelihood

- FOC w.r.t β_1:

$$
\begin{aligned}
\frac{\partial \log L_T(\theta)}{\partial \beta_1} &= \frac{\partial \log L_T(\theta)}{\partial \sigma_t^2}\frac{\partial \sigma_t^2}{\partial \beta_1} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(-\frac{1}{2\sigma_t^2}+\frac{r_t^2}{2\sigma_t^4}\right)\cdot \sigma_{t-1}^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2\sigma_t^2}\left(\frac{r_t^2}{\sigma_t^2}-1\right)\sigma_{t-1}^2
\end{aligned}
$$

The maximum likelihood estimator $\widehat{\theta}_T = (\widehat{\omega}_T, \widehat{\alpha}_{1,T})$ of $\theta = (\omega, \alpha_1)$ can be solved from the equations

$$\begin{aligned} \frac{\partial \log L_T(\widehat{\theta}_T)}{\partial \omega} &= 0, \\ \frac{\partial \log L_T(\widehat{\theta}_T)}{\partial \alpha_1} &= 0, \\ \frac{\partial \log L_T\left(\theta\right)}{\partial \beta_1} &= 0. \end{aligned}$$

GARCH(p,q) model: estimation using R

- When using R, need to install the following package
  - install.packages(TSA)
- R command: garch(x, order = c(p,q))
  - x is the date vector
  - c(p,q) specifies the order of the GARCH model

Estimating GARCH(1,1) with Apple Shares

- We use the daily CC returns of Apple shares from Jan 2, 2015 to Dec 29, 2017 (755 trading dates and hence 754 observations)
- We use the data to fit a GARCH(1,1) model

$$
\begin{aligned}
r_t &= \sigma_t e_t, \;\; e_t \sim iid \;\; N(0,1) \\
\sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\end{aligned}
$$

- We will get the estimators of ω, α_1 and β_1
- By estimating the model, we will also get the residuals $\widehat{e}_t$ and the fitted values $\widehat{\sigma}_t$ for $t = 2, 3, \ldots, 754$
- Let $\widehat{r}_t = \widehat{\sigma}_t \widehat{e}_t$ denote the fitted value of r_t for $t = 2, 3, \ldots, 754$. We can compare r_t and $\widehat{r}_t$ to see the fit of the model

Estimation and Inference of GARCH(1,1)

		ω	α_1	β_1
Apple	est.	0.000020	0.0833	0.8193
	std.	0.000006	0.0220	0.0448

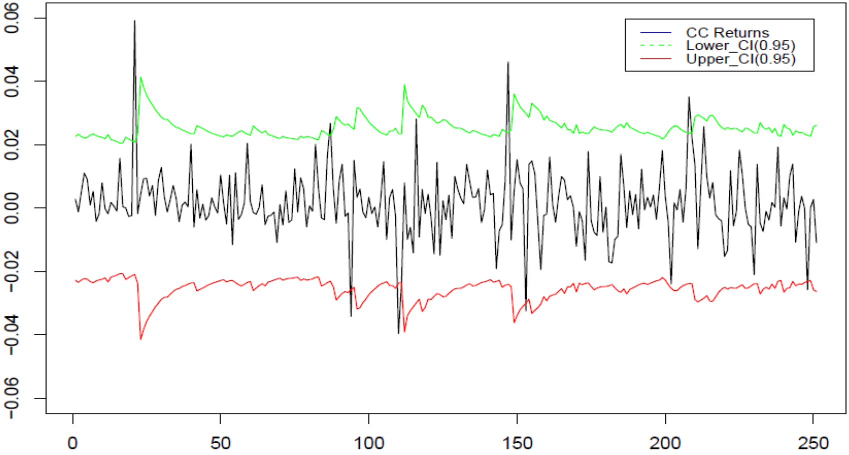
Model: $r_t = \sigma_t e_t$ and $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, $e_t \sim iid\ N(0,1)$

Estimating GARCH(1,1) with Apple Shares

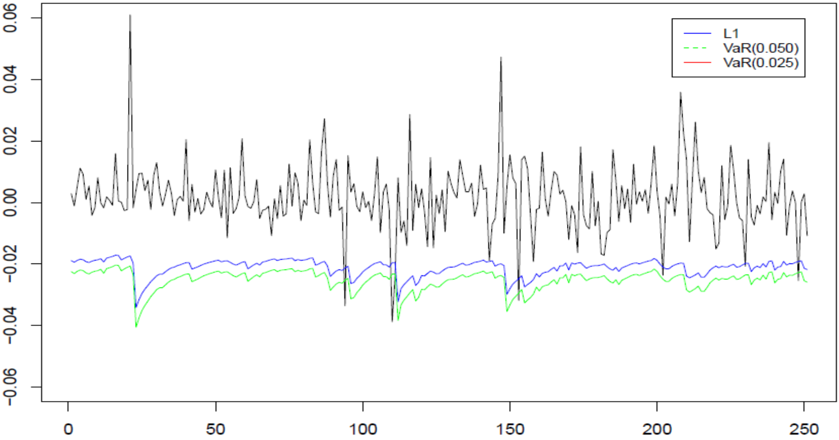
- Empirical results show that the GARCH(1,1) fits the data very well!
- More important question is on the GARCH(1,1)'s performances in prediction
- We are interested in future (tomorrow's) return r_{t+1} given the information up to today
  - building confidence intervals: $\left[\sigma_{t+1} q^Z_{1-\alpha/2},\ \sigma_{t+1} q^Z_{1-\alpha/2}\right]$
  - calculating the VaRs: $W_0 \cdot \left(e^{\sigma_{t+1} q^Z_\alpha} - 1\right)$
- We can use the data up to today to estimate GARCH(1,1) and then get estimator $\widehat{\sigma}_{t+1}$ of σ_{t+1}
  - building confidence intervals: $\left[\widehat{\sigma}_{t+1} q^Z_{1-\alpha/2},\ \widehat{\sigma}_{t+1} q^Z_{1-\alpha/2}\right]$
  - calculating the VaRs: $W_0 \cdot \left(e^{\widehat{\sigma}_{t+1} q^Z_\alpha} - 1\right)$

Estimating GARCH(1,1) with Apple Shares

- We consider the confidence intervals and VaRs for the cc Returns in year 2017
- From Jan 3, 2017, in each trading date of 2017 (there are 253 trading dates), we use 503 observations immediately before this date to estimate a GARCH(1,1)
- We then get an estimator $\widehat{\sigma}_{t+1}$ for this trading date
  - building confidence intervals: $\left[\widehat{\sigma}_{t+1} q^Z_{1-\alpha/2},\ \widehat{\sigma}_{t+1} q^Z_{1-\alpha/2}\right]$
  - calculating the VaRs: $1 \cdot (e^{\widehat{\sigma}_{t+1} q^Z_\alpha} - 1)$
- We are interested in
  - the relative frequency that $\left[\widehat{\sigma}_{t+1} q^Z_{1-\alpha/2},\ \widehat{\sigma}_{t+1} q^Z_{1-\alpha/2}\right]$ covers r_{t+1}
  - the relative frequencies that $L_1 = 1 \cdot (e^{r_{t+1}} - 1)$ exceeds VaR(0.05) and VaR(0.025)



Note: The empirical coverage probability is 0.96414



Note: the relative frequencies that $L_1 = 1 \cdot (e^{r_{t+1}} - 1)$ exceeds VaR(0.05) and VaR(0.025) are 0.0279 and 0.0199 respectively.

Concluding Remarks I

- Some popular volatility models (ARCH/GARCH/SV) are discussed
  - empirical stylized fact, modeling and corresponding properties
  - estimation and inference
  - dominating models in practice
- Next topics
  - portfolio choice and econometrics
  - factor pricing model (estimation and inference)