STAT302: Time Series Analysis

Chapter 11. Time Series Models for Financial Data

Sangbum Choi, Ph.D

Department of Statistics, Korea University

Outline

Asset return

- In finance, we are more interested in the so-called **return** instead of prices of assets.
  - More attractive to investors since it provides simple summary of investment opportunity.
  - Has more attractive (and tractable) statistical features.
- Two popular returns: P_t is a price (or index) of an asset.
  - Simple return

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

  - Log return

$$r_t = \log P_t - \log P_{t-1} = \log(1 + R_t) \approx R_t.$$

Asset Return

- Log-return is more tractable and has some advantages:
  - Multiperiod simple return:

$$\frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}}$$

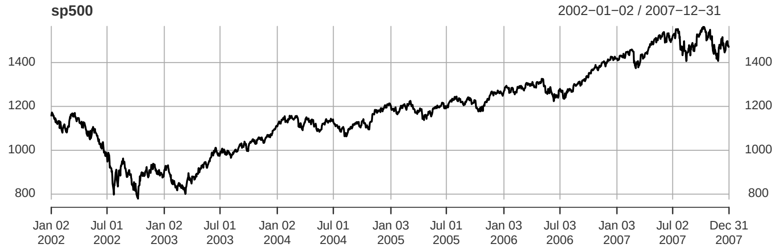
  - Multiperiod log-return:

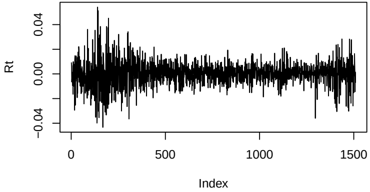
$$\log \frac{P_t}{P_{t-k}} = r_t + r_{t-1} + \cdots + r_{t-k+1}$$

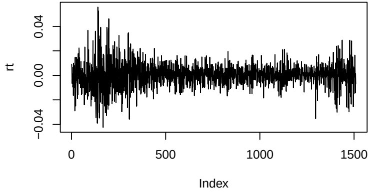
    Hence it is additive and more tractable to see statistical properties.
- In general, log-transformation reduces heteroscedasticity in some sense. Recall Box-Cox transformation or variance stabilization transformation.

S&P 500 stock index

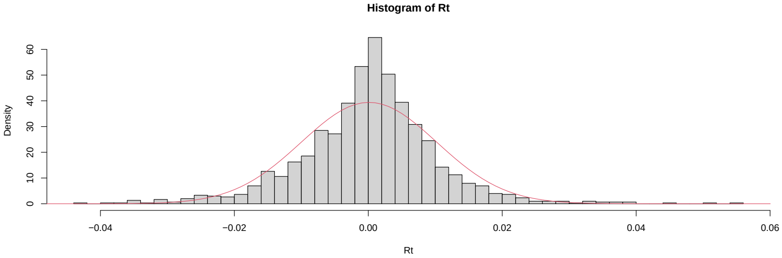
Consider the return of S&P 500 stock index from 2002 to 2007.

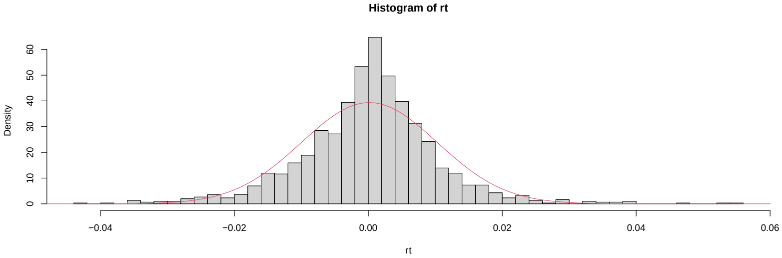






S&P 500 stock index

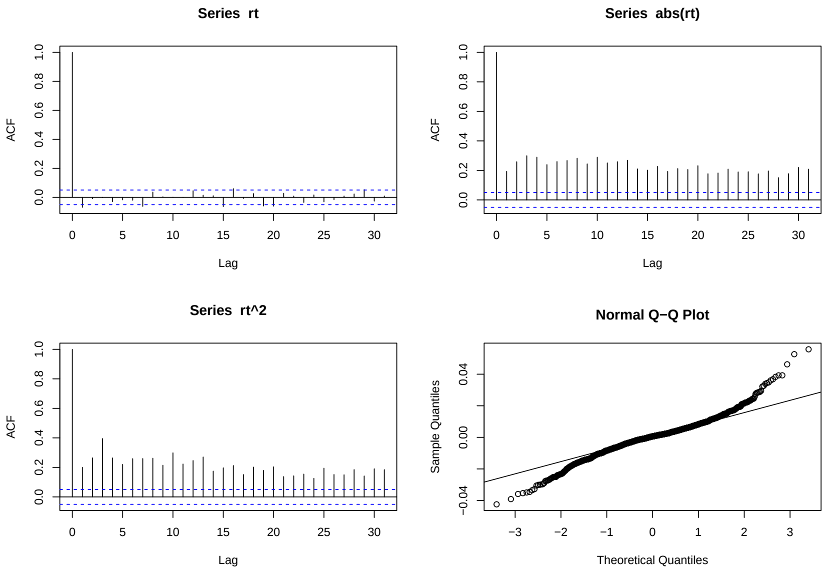




S&P 500 stock index

```
## Normally distributed? No...
ks.test(Rt,"pnorm",0.00011,0.01)
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  Rt
## D = 0.061916, p-value = 1.89e-05
## alternative hypothesis: two-sided
ks.test(rt,"pnorm",0.00016,0.01)
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  rt
## D = 0.060337, p-value = 3.383e-05
## alternative hypothesis: two-sided
```

S&P 500 stock index



Stylized facts about returns

Statistical properties of asset return are

- Symmetric about mean
- r_t itself has little correlations
- However, $|r_t|$ or $|r_t|^2$ has very strong correlations.
- Long-tailed!! Heavy-tailed distribution.
- Variable volatility, that is, **local variance of the process changes across the time (conditional heteroscedasticity)**.

Conditional homoscedasticity

- Information set (or filtration) is the "σ-field" generated by $\xi_{t-1}, \xi_{t-2}, \cdots$,

$$\mathcal{F}_{t-1} = \sigma(\xi_s, s < t) = \sigma(\xi_{t-1}, \xi_{t-2}, \cdots)$$

  It denotes all information available at time $(t-1)$.

- The conditional mean and variance is defined as

$$\mu_t := E(r_t|\mathcal{F}_{t-1}), \quad \sigma_t^2 := \mathrm{Var}(r_t|\mathcal{F}_{t-1})$$

- A process $\{\epsilon_t\}$ is called the **conditional homoscedasticity** iff

$$\mathrm{Var}(\epsilon_t|\mathcal{F}_{t-1}) = constant$$

- Example : $\{\epsilon_t\} \sim$ IID $N(0, 1)$ with $\mathcal{F}_{t-1} = \sigma(\epsilon_s|s < t)$.

Conditional heteroscedasticity

- Suppose that return is a zero mean Gaussian process. To model a varying volatility, consider $r_t \sim N(0,\ \sigma_t^2)$, i.e.,

$$r_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID } N(0,1)$$

  and assume that σ_t is independent of ϵ_t.
- We also observed that $\{r_t^2\}$ has strong correlations, so simply assume an AR(1) structure with multiplicative errors

$$r_t^2 = (\alpha_0 + \alpha_1 r_{t-1}^2\ )\epsilon_t^2, \quad \alpha_0 > 0, \alpha_1 \geq 0$$

- That is, it is assumed that

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

Outline

ARCH(1) model

- A process $\{r_t\}$ is an ARCH(1) model (Autoregressive Conditional Heteroscedasticity) if

$$r_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2.$$

- Technical assumptions: We will assume that σ_t is a measurable function with respect to $\mathcal{F}_{t-1} = \sigma(r_u | u < t)$ and ϵ_t is IID random variable with zero mean and unit variance, being independent of $\mathcal{F}_{t-1}$.

ARCH(1) model

- Then, ARCH(1) is a zero-mean conditional heteroscedasticity model

$$
\begin{aligned}
E(r_t|\mathcal{F}_{t-1}) &= E(\sigma_t\epsilon_t|\mathcal{F}_{t-1}) = \sigma_t \cdot 0 = 0 \\
\text{Var}(r_t|\mathcal{F}_{t-1}) &= E(r_t^2|\mathcal{F}_{t-1}) = E(\sigma_t^2\epsilon_t^2|\mathcal{F}_{t-1}) \\
&= E((\alpha_0 + \alpha_1 r_{t-1}^2)\epsilon_t^2|\mathcal{F}_{t-1}) \\
&= (\alpha_0 + \alpha_1 r_{t-1}^2)E(\epsilon_t^2|\mathcal{F}_{t-1}) \\
&= \alpha_0 + \alpha_1 r_{t-1}^2
\end{aligned}
$$

ARCH(1) model

- ARCH(1) is a nonlinear process.

$$
\begin{aligned}
r_t^2 &= (\alpha_0 + \alpha_1 r_{t-1}^2)\epsilon_t^2 \\
&= \alpha_0 \epsilon_t^2 + \alpha_1 \epsilon_t^2 \{(\alpha_0 + \alpha_1 r_{t-2}^2)\epsilon_{t-1}^2\} \\
&= \alpha_0 \epsilon_t^2 + \alpha_0 \alpha_1 \epsilon_t^2 \epsilon_{t-1}^2 + \alpha_1^2 r_{t-2}^2 \epsilon_t^2 \epsilon_{t-1}^2 \\
&\vdots \\
&= \alpha_0 \left(\sum_{j=0}^{n} \alpha_1^j \epsilon_t^2 \epsilon_{t-1}^2 \cdots \epsilon_{t-j}^2 \right) + \alpha_1^{n+1} r_{t-n-1}^2 \epsilon_t^2 \epsilon_{t-1}^2 \cdots \epsilon_{t-n}^2
\end{aligned}
$$

Stationarity of ARCH(1)

If $0 \le \alpha_1 < 1$, then

(1) $\{r_t\}$ is stationary since it is a function of IID $\{\epsilon_s, s < t\}$ and the last term has expectation

$$\alpha_1^{n+1} E(r_{t-n-1}^2) \to 0$$

(2) $E(r_t) = E(E(r_t|\mathcal{F}_{t-1})) = 0$

(3) $E(r_{t-1}^2) = \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j = \frac{\alpha_0}{1-\alpha_1}$

(4) $\mathrm{Cov}(r_{t+h}, r_t) = E(r_{t+h} r_t) = E(E(r_{t+h} r_t|\mathcal{F}_{t+h-1})) = E(r_t)E(r_{t+h}|\mathcal{F}_{t+h-1}) = 0$. Therefore, $\mathrm{Cor}(h) = 0$ if $h \neq 0$ (r_t has little correlations)

(2)–(4) imply that $\{r_t\}$ is a WN sequence.

Solution of ARCH(1)

- If $0 \leq \alpha_1 < 1$, the unique causal stationary solution of the ARCH(1) equations is given by

$$r_t = \epsilon_t \sqrt{\alpha_0 \left(\sum_{j=0}^{n} \alpha_1^j \epsilon_t^2 \epsilon_{t-1}^2 \cdots \epsilon_{t-j}^2 \right)}$$

- Recall that $\{r_t\}$ is a WN sequence, but not IID sequence. Note that

$$E(r_t^2 | r_{t-1}) = (\alpha_0 + \alpha_1 r_{t-1}^2) E(\epsilon_t^2 | r_{t-1}) = \alpha_0 + \alpha_1 r_{t-1}^2.$$

ARCH(1) properties: persistency

- Observe ARCH(1) dynamics

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

- If r_{t-1}^2 is large (small), then σ_t^2 also become large (small). Then, it leads to large (small) r_t^2 due to relationship $r_t^2 = \sigma_t^2 \epsilon_t^2$. Finally it will also produce larger (smaller) r_{t+1}^2.
- It means that once larger value of r_t^2 is observed, then it tends to produce larger r_{t+1}^2.

ARCH(1) properties: persistency

- This is what is observed for ACF of r_t^2 - **persistency of volitility!**
- One can show that for all $h \in \mathbb{Z}$,

$$\mathrm{Cor}_{r_t^2}(h) = \alpha_1^{|h|},$$

so we expect very strong correlations for $\{r_t^2\}$

ARCH(1) properties: long-tailed

- Under 4th-order condition, $E(r_t^4) = m_4 < \infty$ (constant), one can show that

$$m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$$

- Hence, the Kurtosis

$$\frac{E(r_t^4)}{\mathrm{Var}^2(r_t)} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3$$

- If $1 - 3\alpha_1^2 > 0$, i.e., $0 \le \alpha_1^2 < 1/3$, $\{r_t\}$ has heavier tail than Gaussian. (fourth moment of $N(0,1)$ is 3).
- Furthermore, we can show that

$$E(r_t^{2k}) = +\infty,$$

for some positive k.

Weakness of ARCH models

- The distribution of $\{r_t\}$ symmetric, i.e, $r_t \overset{d}{=} -r_t$ since ϵ_t is symmetric. However, we model r_t^2, so positive & negative shocks has the same effect in model. But, in practice, financial market response differently. It is known as **leverage effect**.
- Too restrictive for moment condition. Finite 4th moment is only possible when $\alpha_1^2 < 1/3$ being $0 \leq \alpha_1 < 0.57$. Not flexible in model estimation.
- May not suitable to explain long-term correlations. This is related to long-range dependence.

ARCH(m)

- ARCH(m) model assumes that

$$r_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2$$

- Observe that

$$r_t^2 \approx \sigma_t^2$$

since $E(\epsilon_t^2) = 1$. Hence

$$r_t^2 \approx \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2$$

ARCH(m)

- We can regrad $\{r_t^2\}$ as AR(m) model

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2 + \eta_t$$

since approximation error $\eta_t = r_t^2 - \sigma_t^2$ is WN

i) $E(\eta_t) = E(E(\eta_t|\mathcal{F}_{t-1})) = E(E(r_t^2 - \sigma_t^2|\mathcal{F}_{t-1})) = E(E(\sigma_t^2\epsilon_t^2 - \sigma_t^2|\mathcal{F}_{t-1})) = 0$, since ϵ_t is independent of $\mathcal{F}_{t-1}$.

ii) Similarly, $E(\eta_t\eta_{t-1}) = 0$ (uncorrelated)

Testing for ARCH effect

- Interested in formal testing of conditional heteroscedasticity, known as **ARCH effect**.
- Recall that we can consider $\{r_t^2\}$ be AR(m) process.
- First test is simply use Ljung-Box statistics $Q(m)$ to $\{r_t^2\}$.
- The second test is called Engle's ARCH test based on Lagrange Multiplier.

Testing for ARCH effect

- For

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2 + \eta_t, \quad t = m+1, \cdots, T,$$

test whether $\alpha_1 = \alpha_2 = \cdots = \alpha_m$ from

$$F = \frac{(SST - SSR)/m}{SSR/(T - m - m - 1)} \approx \chi^2(m),$$

where $SST = \sum_t (r_t^2 - \bar{r_t^2})^2$ and $SSR = \sum_t \hat{\eta}^2$.

Estimation of ARCH(m)

- Conditional MLE: By assuming Gaussianity, we have

$$
\begin{aligned}
f(r_1, r_2, \cdots, r_t | \alpha_0, \cdots, \alpha_m) & \\
&= f(r_t|\mathcal{F}_{t-1}) f(r_{t-1}|\mathcal{F}_{t-2}) \cdots f(r_2|\mathcal{F}_1) f(r_1) \\
&\approx \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp(-\frac{r_t^2}{2\sigma_t^2})
\end{aligned}
$$

- Hence, truncating few observations yield conditional MLE as

$$
\textit{minimize} \sum_{t=m+1}^{T} (\log\ \sigma_t^2 + \frac{r_t^2}{\sigma_t^2})
$$

- Maybe use the quasi-Maximum likelihood for GARCH(p, q).

Model checking and forecasting

- From the relationship, $r_t = \sigma_t \epsilon_t$, we can estimate innovations

$$\hat{\epsilon}_t = \frac{r_t}{\hat{\sigma}_t} \approx \mathrm{IID}(0, \sigma^2)$$

- Standard tools such as Ljung-Box, QQ-plot, residual plots to check indeed they are IID sequence.
- Forecasting: Best Linear Unbiased Predictor (BLUP). Apply iteratively the following:

$$\begin{aligned}
\hat{\sigma}^2_{t+1} &= \hat{\alpha}_0 + \hat{\alpha}_1 r_t^2 + \cdots + \hat{\alpha}_m r^2_{t-m} \\
\hat{\sigma}^2_{t+2} &= \hat{\alpha}_0 + \hat{\alpha}_1 (\hat{\alpha}^2_{t+1}) + \hat{\alpha}_2 r_t^2 + \cdots + \hat{\alpha}_m r^2_{t-m+1} \\
&\vdots
\end{aligned}$$

- However, forecasting r^2_{t+h} is not common due to multiplicative error though $r^2_{t+h} \approx \sigma^2_{t+h}$.

GARCH(p, q) models

- Extension of ARCH(m) model by including moving average part.
- GARCH(p, q) models :

$$r_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, \sigma^2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

- Stationarity condition : $\alpha_0 > 0, \alpha_j, \beta_j \geq 0, \sum \alpha_i + \sum \beta_j < 1.$

GARCH(p, q) models

- GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

- Using $r_t^2 \approx \sigma_t^2$ and $\eta_t = r_t^2 - \sigma_t^2$ implies

$$r_t^2 = \alpha_0 + (\alpha_1 + \beta_1) r_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}.$$

- Hence, $\{r_t^2\}$ is an ARMA(1,1) model.

Outline

ARMA-GARCH model

- ARCH handles conditional (hetero) variance, what about conditional mean?
- Consider AR(1) model:

$$u_t = \phi u_{t-1} + \epsilon_t$$

- Then, it is well known that

$$E(u_t) = 0, \quad \text{Var}(u_t) = \frac{\sigma^2}{1-\phi^2}$$

- Hence, **unconditional mean and variance are constant**.

ARMA-GARCH model

- Now, consider conditional mean/variance

$$E(u_t|\mathcal{F}_{t-1}) = E(\phi u_{t-1} + \epsilon_t|\mathcal{F}_{t-1}) = \phi u_{t-1}$$

  Thereby, **conditional mean is non-constant!**

- Also, conditional variance is given by

$$E(u_t^2|\mathcal{F}_{t-1}) = E(\phi^2 u_{t-1}^2 + 2\epsilon_t \phi u_{t-1} + \epsilon_t^2|\mathcal{F}_{t-1}) = \phi^2 u_{t-1}^2 + \sigma^2$$

- So, $\text{Var}(u_t|\mathcal{F}_{t-1}) = E(u_t^2|\mathcal{F}_{t-1}) - E(u_t|\mathcal{F}_{t-1})^2 = \sigma^2$.

ARMA-GARCH model

- Notice that

	ARMA	GARCH
marginal mean	constant	constant
marginal variance	constant	constant
conditional mean	non-constant	constant
conditional variance	constant	non-constant

- Therefore, ARMA-GARCH model incorporates non-constant conditional mean and variance

$$\begin{cases} X_t = \sum_{i=1}^{P} \phi_i X_{t-i} + \sum_{j=1}^{Q} \theta_j \epsilon_{t-j} \\ \epsilon_t = \sigma_t \xi_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \end{cases}$$

ARCH testing

- You can use `ArchTest()` function in `FinTS` library to check the existence of heterogeneity in variance.
- Let $\hat{r}_t$ denote the estimated residual process by ARCH(p). Then, the ARCH test is testing

$$H_0 : \hat{r}_t^2 = \alpha_0 \quad \text{v.s.} \quad H_1 : \hat{r}_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \hat{r}_{t-j}^2$$

- That is, unless $\alpha_1 = ... = \alpha_p$, we can conclude that there exists heterogeneity in the variance of the return.

ARCH testing

```
library(tidyverse)
library(xts)
library(FinTS)
d=gt::sp500 %>% select(date,adj_close)
sp500=xts(d$adj_close,d$date)
sp500=window(sp500,
             start=as.Date("2002-1-1"),
             end=as.Date("2007-12-31")) %>% as.vector
rt=diff(log(sp500))
sapply(1:10, function(i) ArchTest(rt,lags=i)$p.value) %>% round(3)
## Chi-squared Chi-squared Chi-squared Chi-squared Chi-squared Chi-squared
##           0           0           0           0           0           0
## Chi-squared Chi-squared Chi-squared Chi-squared
##           0           0           0           0
```

GARCH modeling

```
auto.arima(rt,stationary=TRUE)
## Series: rt
## ARIMA(0,0,1) with zero mean
##
## Coefficients:
##           ma1
##       -0.0702
## s.e.   0.0259
##
## sigma^2 = 0.0001022:  log likelihood = 4792.21
## AIC=-9580.41   AICc=-9580.41   BIC=-9569.78
```

GARCH modeling

```
library(fGarch)
garchFit(data=rt,formula=~arma(0,1) ~garch(1,1), trace=FALSE)
##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ~arma(0, 1) ~ garch(1, 1), data = rt, trace = FALSE)
##
## Mean and Variance Equation:
##  data ~ garch(1, 1)
## <environment: 0x11012cd30>
##  [data = rt]
##
## Conditional Distribution:
##  norm
##
## Coefficient(s):
##         mu       omega      alpha1       beta1
## 3.9694e-04  8.8320e-07  5.3166e-02  9.3631e-01
##
## Std. Errors:
##  based on Hessian
##
```

GARCH Zoo

- There are lots of variations in GARCH to overcome shortcomings of GARCH models such as leverage effect, non-stationarity, long-term correlations etc. It includes
  - AGARCH: (Asymmetric) GARCH
  - MGARCH: (Multivariate) GARCH
  - TARCH: Threshold ARCH
  - NGARCH: Nonlinar GARCH
  - EGARCH: Exponential GARCH
  - IGARCH: Integrated GARCH
- See Tsay (2004) Ch3.6–3.14. Here we briefly introduce EGARCH and IGARCH.

Exponential GARCH and leverage effect

- Want to account for asymmetric effects between positive and negative asset returns. Also, we expect that negative shocks tend to have larger impacts.
- EGARCH(m, s) is given by

$$r_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim \mathrm{IID}(0, \sigma^2)$$

$$\log(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + \cdots + \beta_{s-1} B^{s-1}}{1 - \alpha_1 B - \cdots - \alpha_m B^m} g(\epsilon_{t-1}),$$

where,

$$g(\epsilon_t) = \begin{cases} (\theta + \gamma)\epsilon_t - \gamma E(|\epsilon_t|) \\ (\theta - \gamma)\epsilon_t - \gamma E(|\epsilon_t|) \end{cases}$$

Exponential GARCH and leverage effect

- Took log for positive value.
- For example EGARCH(1,1) is given by

$$(1-\alpha_1 B)\log(\sigma_t^2) = \begin{cases} \alpha_* + (\theta+\gamma)\epsilon_{t-1} & \epsilon_{t-1} \geq 0 \\ \alpha_* + (\theta-\gamma)(-\epsilon_{t-1}) & \epsilon_{t-1} < 0 \end{cases}$$

- Depending on the sign of innovations, it has different dynamics for dependence.
- It is also written as

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{s} \alpha_i \frac{|r_{t-i}| + \gamma_i r_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{m} \beta_j \log(\sigma_{t-j}^2)$$

and γ_i controls leverage effect. (We hope this to be negative in real application)

More on long-term correlations

- IGARCH assumes that

$$\sum_i \alpha_i + \sum_j \beta_j = 1$$

- However, it has the following issues:
  - Neither stationary nor non-stationary since it has "infinite" variance.
  - Infinite variance means that it is heavy-tailed. People are skeptical about this point.
- Alternative modeling is time-varying GARCH model, Threshold GARCH model etc if you keep GARCH-type of structure in the model.
- Another approach is called the long-range dependence time series.