LECTURE 9

5 Conditional Heteroscedastic Models

The models referred to as conditional heteroscedastic models are used for modeling the volatility of asset returns. The volatility means the conditional standard deviation of the underlying asset return. It plays important role in

- calculating value at risk
- asset allocation
- helps to improve the efficiency in parameter estimation and the accuracy in interval forecast.

The volatility index VIX is a financial instrument traded in financial markets.

— ViX represents markets expectations for volatility over the coming 30 days; VIX is used to measure the level of risk or stress in the market when making investment decisions.

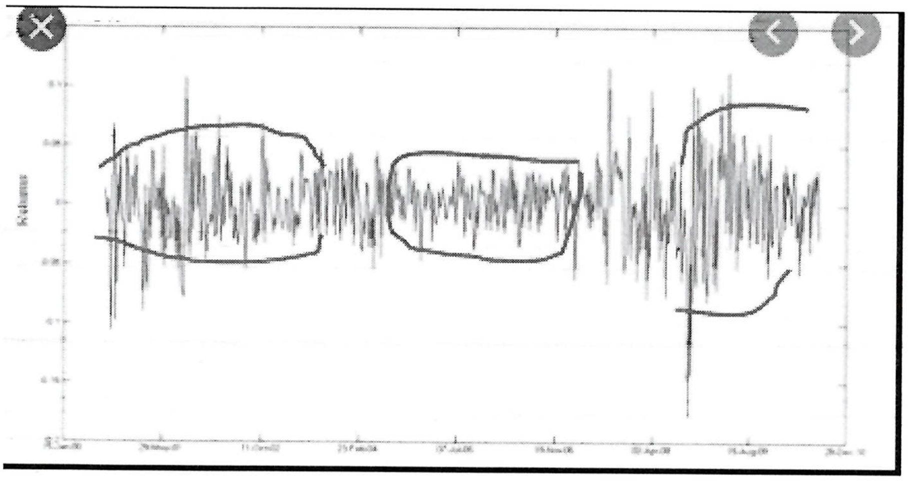
We shall discuss:
– the autoregressive conditional heteroscedastic model (ARCH) of Engle (1982)
– the generalized ARCH model (GARCH) by Bollerslev (1982) and show some applications.

5.1 Characteristics of Volatility

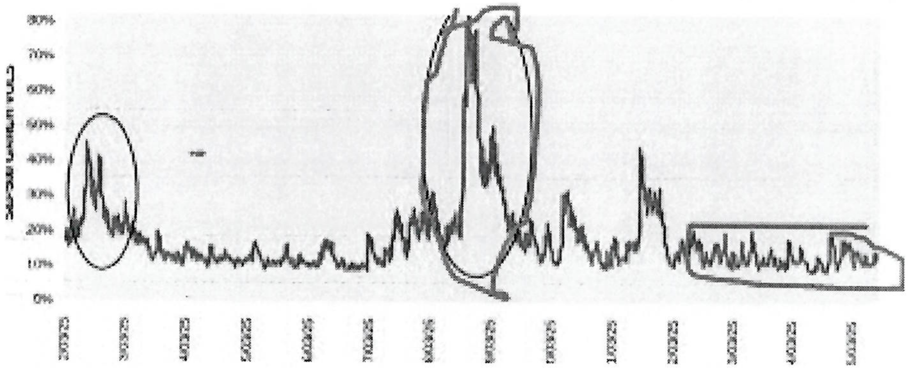
The feature of volatility is that volatility is not directly observable. It has some characteristics that are commonly seen in asset returns:

- There exist volatility clusters (volatility might be high for certain time periods and low for other periods)

  "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes" Mandelbrot
- Volatility evolves in continuous manner, volatility jumps are rare.
- Volatility does not diverge to infinity, ir remains within some fixed range. Statistically speaking, volatility is often stationary.
- It reacts differently to a big price increase or a big price drop. This effect is called leverage effect.



Volatility Clustering and Agent Based ...
zanaducloud.com



S&P 500 VOLATILITY CLUSTERING 90 ...
researchgate.net

These properties play an important role in developing volatility models, which are trying to capture mentioned characteristics.

5.1.1 Structure of the models

Let r_t be the log return of an asset P_t at time t:

$$r_t = \log P_t - \log P_{t-1}$$

Basic idea behind volatility study:

— series of log returns r_t is uncorrelated or with minor lower order correlation,

— but series of log returns r_t is dependent series.

Example. Figure 3.4 shows the ACF of the monthly stock returns of Intel corporation for 1973 to 2003:

- Figure 3.4 a) shows the sample ACF of the returns r_t which suggests no significant correlation
- Figure 3.4 c) shows the sample ACF of the absolute log return $|r_t|$ which is correlated
- Figure 3.4 b) shows the sample ACF of the squared returns r_t^2 which suggest correlation

Conclusion: Figures show that monthly returns r_t are serially uncorrelated, but dependent. Volatility models try to capture such dependence.

Notice: if e_t are independent variables then
e_t^2, $|e_t|$ are also independent variables.
Independent variables are uncorrelated. Therefore all three series
e_t e_t^2, $|e_t|$ are uncorrelated.

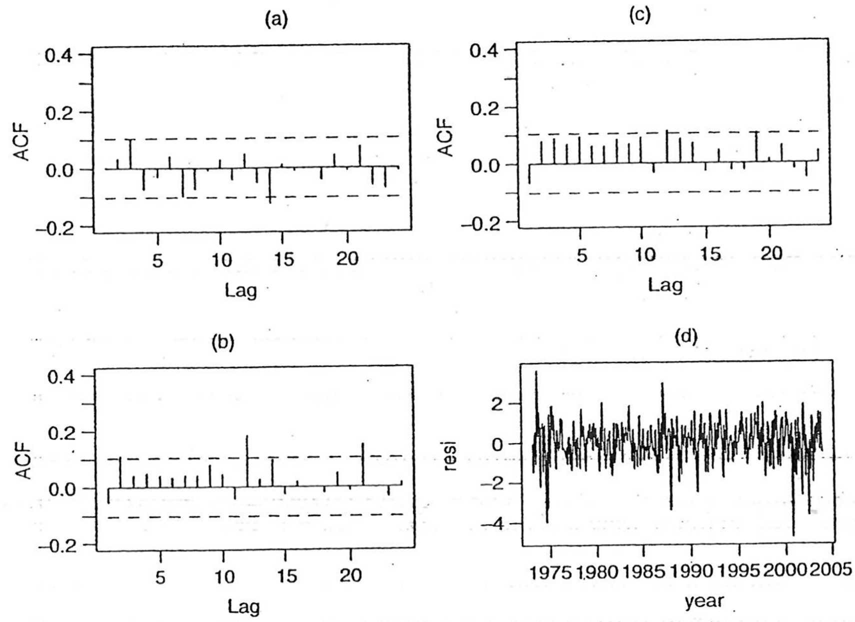
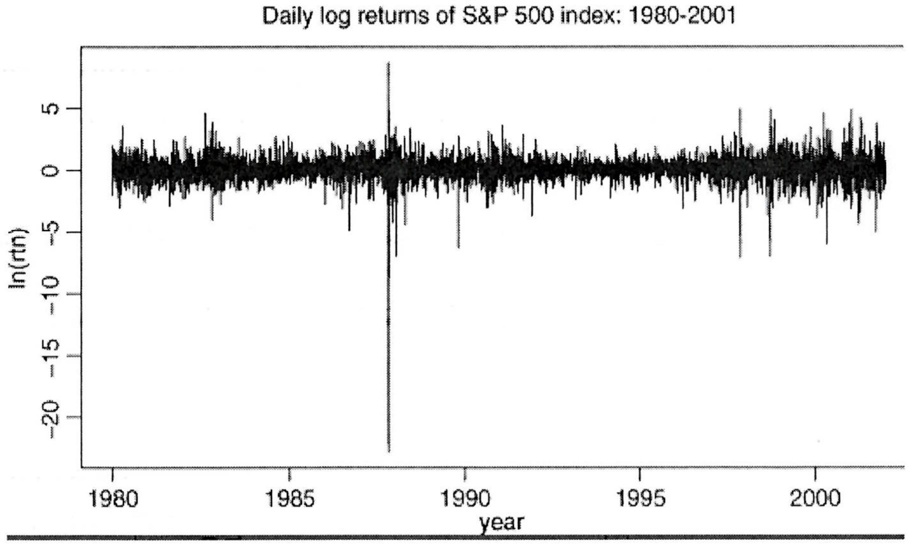
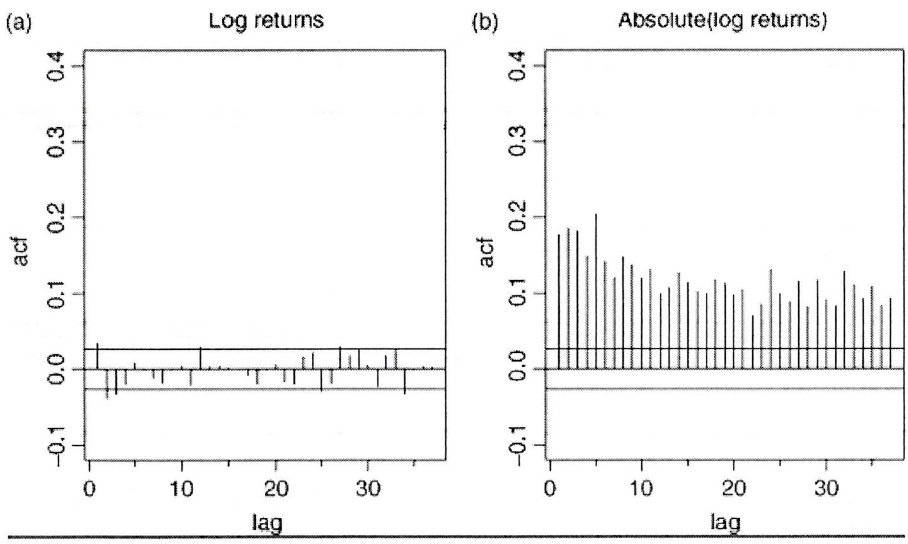


Figure 3.4. Model checking statistics of the Gaussian ARCH(1) model in Eq. (3.11) for the monthly log returns of Intel stock from January 1973 to December 2003: parts (a), (b), and (c) show the sample ACF of the standardized residuals, their squared series, and absolute series, respectively; part (d) is the time plot of standardized residuals.

Example. Plot of daily log returns r_t of $S\$P$ index

Sample ACF of r_t does not show correlation in r_t.
Sample ACF of $|r_t|$ shows strong correlation in $|r_t|$.
Sample ACF of r_t^2 would also show strong correlation in r_t^2.





Definition of volatility.

We introduce

– conditional mean μ_t

– conditional variance σ_t^2

given the information F_{t-1} available at time $t-1$:

$$\mu_t = E(r_t|F_{t-1}), \qquad \sigma_t^2 = Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}].$$

We observed in previous example that serial correlation in returns r_t is weak.

First step: Modeling μ_t. The equation for the mean μ_t should be simple. Often we set $\mu_t = \mu$ constant!

If r_t are correlated, μ_t can be modeled by and ARMA(p,q) model with some explanatory variables $\xi_{i,t}$:

$$r_t = \mu_t + r_t^*, \quad \mu_t = \phi_0 + \sum_{i=1}^{p} \beta_i r_{t-i} - \sum_{i=1}^{q} \theta_i {r^*}_{t-i} + \sum_{i=1}^{p} \phi_i \xi_{i,t-i},$$

where

– $\xi_{i,t}$ are some explanatory variables,

– p and q is the order of an ARMA model and

– r_i^* are uncorrelated variables (white noise) with 0 mean.

The order p, q might depend on frequency of the data:

— daily returns r_t might show some minor correlation,

— monthly returns r_t tend to be uncorrelated. Then we set $\mu_t = \mu$.

A dummy variable $\xi_{i,t}$ might be for Mondays to study weekend effect.

Combining equations we have

$$\sigma_t^2 = Var(r_t|F_{t-1}) = Var(r_t^*|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}].$$

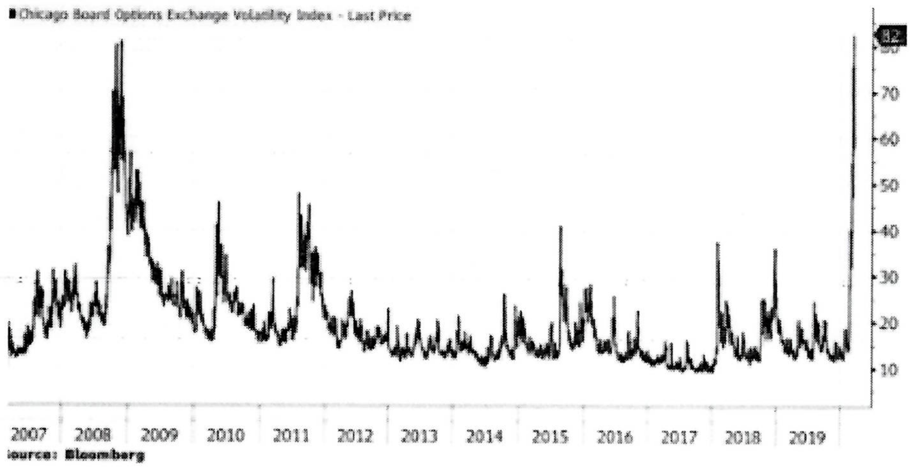
Definition: σ_t^2 is called volatility.

We will model evolution of the volatility σ_t^2 in time. Such evolution will be governed by

— ARCH -GARCH models.

Besides these models volatility can be modeled by stochastic volatility models where σ_t^2 is defined by a stochastic equation.

Example Volatility plot. Note: volatility σ_t^2 in non-negative, it is not observed directly! We need to estimate/ extract it from data.



5.1.2 Model building

Constructing volatility model for a return series consists of three steps:

- Specify a model for a mean μ_t. If no correlation in returns r_t, one can set

$$\mu_t = \mu$$

  constant. If returns r_t correlated, you can fit a model to r_t. It can be an ARMA model, if needed, to remove dependence.

- Use residuals $r_t^* = r_t - \mu_t$, to test for ARCH effects in r_t^*.

  ARCH effect means:

  – r_t^* are uncorrelated,

  – the squares r_t^* correlated

- If ARCH effect is significant, specify volatility model for σ_t^2 and perform joint estimation of μ_t and σ_t^2.

For most asset return series r_t, the correlation is weak. So mean equation $r_t = \mu + r_t^*$ results in removing the sample mean $\bar{r}$ from the data r_t if the sample mean is significantly different from zero:

$$r_t^* = r_t - \bar{r}, \quad \bar{r} = T^{-1} \sum_{j=1}^{T} r_j.$$

For daily returns, a simple AR(p) model might be needed to be fitted to r_t to remove correlation in r_t.

The mean equation may also employ some explanatory variables, such as dummy variables for weekend or January effects.

5.2 Testing for ARCH effect

ARCH effect. Testing for ARCH effect (or conditional heteroscedasticity) in the series $r_t = \mu_t + r_t^*$, means testing for

- no correlation in residuals $r_t^* = r_t - \mu_t$
- correlation in squares r_t^*

Two tests can be used.

Ljung-Box test. We can apply Ljung-Box statistics $Q(m)$ to the squares r_t^{*2}. Then the null hypothesis is that the first m ACF's are zero:

$$\rho_1 = 0, \quad \rho_2 = 0, \quad ..., \quad \rho_m = 0.$$

The Lagrange multiplier test (by Engle (1982)). It is the usual test F test testing for the null hypothesis

$$H_0 : \quad \alpha_1 = ... = \alpha_m = 0$$

in the linear regression

$$r_t^{*2} = \alpha_0 + \alpha_1 r_{t-1}^{*2} + ... + \alpha_m r_{t-m}^{*2} + e_t, \quad t = m+1, ..., N$$

where e_t denotes the error term, $m \geq 1$ is the lag and N is the sample size.

Note: Under H_0,

$$r_t^{*2} = \alpha_0 + e_t$$

is white noise. Then, r_t^{*2} are uncorrelated variables.

Denote

$$SSR_0 = \sum_{t=m+1}^{N} (r_t^{*2} - \bar{\omega})^2, \qquad where \quad \bar{\omega} = (1/N)\sum_{t=1}^{N} r_t^{*2},$$

$$SSR_1 = \sum_{t=m+1}^{N} \hat{e}_t^2, \qquad \hat{e}_t = r_t^{*2} - \hat{\alpha}_0 - \hat{\alpha}_1 r_{t-1}^{*2} - \ldots - \hat{\alpha}_m r_{t-m}^{*2}$$

where $\hat{e}_t$ is the least squares residual of the previous linear regression.

Then, under H_0, the statistic

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(N - 2m - 1)} \sim \chi_m^2$$

has asymptotical chi-squared distribution with with m degrees of freedom.

Example. Consider the monthly log stock returns of Intel Corporation from 1873 to 2003. The series has no significant correlation so it can be used directly to test for an ARCH effect.

— $Q(12) = 18.57$ with p-value 0.10 confirms no serial correlation in r_t^{*2}. So, no ARCH effect.

– On the other hand, the Lagrange multiplier test shows strong ARCH effect with test statistic $F \sim 43.5$ and p- value close to zero.

S-Plus Demonstration

Denote the return series by `intc`. Note that the command `archTest` applies directly to the a_t series, not to a_t^2.

```
> autocorTest(intc,lag=12)
Test for Autocorrelation: Ljung-Box
Null Hypothesis: no autocorrelation

Test Statistics:
Test Stat 18.5664   p.value   0.0995

Dist. under Null: chi-square with 12 degrees of freedom
Total Observ.: 372

> archTest(intc, lag=12)
Test for ARCH Effects: LM Test
Null Hypothesis: no ARCH effects

Test Statistics:
Test Stat 43.5041   p.value   0.0000

Dist. under Null: chi-square with 12 degrees of freedom
```

5.3 The ARCH model

ARCH model was suggested by Engle (1982). It provides a framework for volatility modeling. The idea of such modeling is based on two facts:

a) the shock r_t^* of asset return is serially uncorrelated, but dependent,

b) the dependence of r_t^* can be described by a simple quadratic function for r_t^{*2}.

ARCH(m) model: it assumes that

$$r_t^* = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^{*2} + ... + \alpha_m r_{t-m}^{*2},$$

where ε_t is a sequence of iid variables with $E\varepsilon_t = 0$, $E\varepsilon_t^2 = 1$,

$$\alpha_0 > 0, \quad \alpha_1 \geq 0, ..., \alpha_m \geq 0.$$

For existence of a stationary solution, the parameter must satisfy additional conditions. Here,

$$\mu_t = E[r_t^* | F_{t-1}] = 0, \qquad \text{var}(r_t^{*2} | F_{t-1}) = \sigma_t^2.$$

In practice, ε_t are often assumed to follow the standard normal $N(0,1)$ or Student t-distribution. In case of this model, the past large shocks $r_{t-1}^{*2}, .., r_{t-m}^{*2}$ implies large conditional variance σ_t^2 for the innovation r_t^*. In ARCH models, large shocks tend to be followed by another large shock. "Tend" means that the probability to obtain large variate is greater than that of smaller variate. So ARCH model has the feature of clustering observed in asset returns.

Example. The ARCH effect is rather common in financial time series. Figure 3.2 shows the time plot of

- the percentage change in Deutche mark/US dollar exchange rate measured in 10 min intervals from June 5, 1989 to June 9, 1989, 2488 observations.
- the squared series of the percentage changes.

We observe, that big changes occur rarely, and there are certain stable periods. Figure 3.3 a) shows that there is no serial correlation in series of percentage changes. Figure 3.3 b) of PACF shows that correlation is present in the squared series of changes. PACF has big spikes, suggesting that series of percentage changes is not independent.

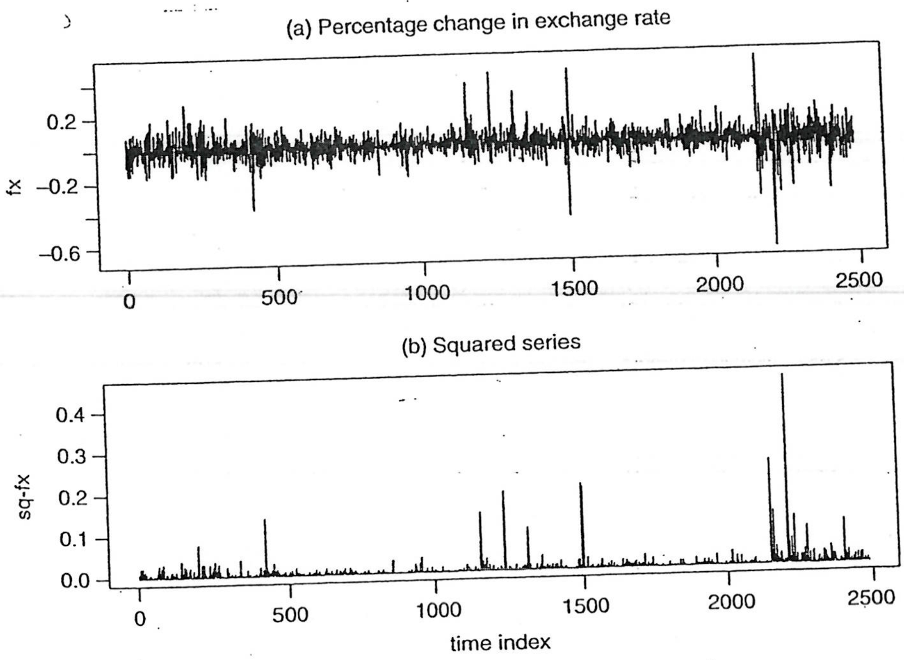


Figure 3.2. (a) Time plot of 10-minute returns of the exchange rate between Deutsche mark and U.S. dollar and (b) the squared returns.

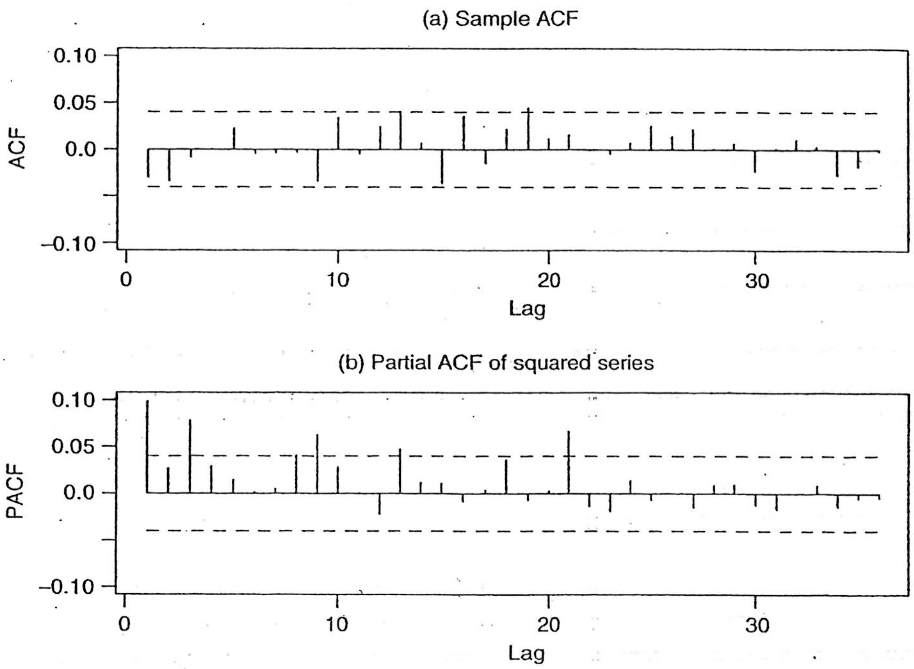


Figure 3.3. (a) Sample autocorrelation function of the return series of mark/dollar exchange rate and (b) sample partial autocorrelation function of the squared returns.

5.4 Properties of $ARCH(1)$ model

ARCH(1) model.
To understand ARCH models, first we study $ARCH(1)$ model

$$r_t^* = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^{*2},$$

with parameters $\alpha_0 > 0, \alpha_1 \geq 0$.

Unconditional mean:

$$\begin{aligned} E[r_t^*] &= E[\sigma_t \varepsilon_t] = E[E[\sigma_t \varepsilon_t | F_{t-1}]] \\ &= E[\sigma_t E[\varepsilon_t | F_{t-1}]] = E[\sigma_t \times 0] = 0. \end{aligned}$$

Unconditional variance:
Notice that: $E[\varepsilon_t^2 | F_{t-1}] = E[\varepsilon_t^2] = 1$. Then,

$$\begin{aligned} Var(r_t^*) &= E[r_t^{*2}] = E[\sigma_t^2 \varepsilon_t^2] = E[E[\sigma_t^2 \varepsilon_t^2 | F_{t-1}]] = E[\sigma_t^2 E[\varepsilon_t^2 | F_{t-1}]] \\ &= E[\sigma_t^2] = E[\alpha_0 + \alpha_1 r_{t-1}^{*2}] = \alpha_0 + E[r_{t-1}^{*2}]. \end{aligned}$$

Because r_t^* is a stationary process, and $E[r_t^*] = 0$, then

$$Var(r_t^*) = Var(r_{t-1}^*) = E[r_{t-1}^{*2}].$$

Hence

$$Var(r_t^*) = \alpha_0 + \alpha_1 Var(r_t^*)$$

and

$$Var(r_t^*) = \frac{\alpha_0}{1 - \alpha_1}.$$

Since the variance must be positive, we require

$$0 \leq \alpha_1 < 1.$$

Tail behaviour. To study tail behaviour, we require finite forth moment of r_t^*. Under assumption of normality of ε_t,

$$E[\varepsilon_t^4 | F_{t-1}] = E[\varepsilon_t^4] = 3.$$

Therefore

$$E[r_t^{*4}] = E[\sigma_t^4\varepsilon_t^4] = E[E(\sigma_t^4\varepsilon_t^4|F_{t-1})] = E[\sigma_t^4 E(\varepsilon_t^4|F_{t-1})] = 3E[\sigma_t^4]$$

$$= 3E(\alpha_0 + \alpha_1 r_{t-1}^{*2})^2 = 3E(\alpha_0^2 + 2\alpha_0\alpha_1 r_{t-1}^{*2} + \alpha_1^2 r_{t-1}^{*4}).$$

If r_t^* is fourth order stationary, then setting $m_4 = E[r_t^{*4}]$ we obtain

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

Therefore

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

Consequences of this result.

- Since m_4 is positive, then α_1 must satisfy condition
$$1 - 3\alpha_1^2 > 0,$$
that is $0 \le a_1^2 < 1/3$,
- the kurtosis of r_t^* is
$$\frac{E(r_t^{*4})}{[Var(r_t^*)]^2} = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}\frac{(1-\alpha_1)^2}{\alpha_0^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3.$$

It is positive and greater than 3.

We see that:

— Excess kurtosis is positive and tail of distribution is heavier that of normal distribution. Since tail is heavy, the model will produce large values (outliers).

— Similar properties continue to hold for general ARCH(p) models.

Weaknesses of ARCH models:

1. ARCH models assumes that positive and negative shocks have the same effect on volatility, what is not observed in practise,
2. The ARCH models impose restrictive assumptions on parameters.
3. ARCH models do not explain source of variation.
4. ARCH models are over-predicting the volatility

5.4.1 Building ARCH model

Specifying the order of ARCH model is rather easy.

Order determination. One can use PACF of r_t^{*2}. The model we have is

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^{*2} + ...\alpha_p r_{t-p}^{*2}.$$

For a given sample,

$$r_t^{*2} \text{ is an unbiased estimator for } \sigma_t^2.$$

Therefore we might expect r_t^{*2} to be linearly related to $r_{t-1}^{*2}, ..., r_{t-p}^{*2}$ in a manner similar to a linear regression. In general, r_t^{*2} is not an efficient estimator for σ_t^2, but an approximation, when we identifying order p.

Alternatively, define

$$\eta_t = r_t^{*2} - \sigma_t^2.$$

It can be shown that η_t is uncorrelated series (white noise) and has zero mean. The ARCH model we can write as an AR(p) model

$$\begin{aligned} r_t^{*2} &= \sigma_t^2 + \eta_t, \\ r_t^{*2} &= \alpha_0 + \alpha_1 r_{t-1}^{*2} + ...\alpha_p r_{t-p}^{*2} + \eta_t, \end{aligned}$$

where η_t is a white noise but not an iid.
The PACF function might be useful determining th order of the model.

Parameter estimation. Parameters of ARCH model can be estimated using (conditional) Maximum Likelihood method both in case when the noise ε_t has normal of Student-t distribution.

Forecasting. Forecast of an ARCH model can be obtained recursively as in case of AR models.

Let be given ARCH(p) model. Then the 1-step ahead forecast of volatility is

$$\begin{aligned} \sigma_t^2(1) &= E[\sigma_{t+1}^2 | F_t] = E[\alpha_0 + \alpha_1 r_t^{*2} + ... + \alpha_p r_{t+1-p}^{*2} | F_t] \\ &= \alpha_0 + \alpha_1 r_t^{*2} + ... + \alpha_p r_{t+1-p}^{*2}. \end{aligned}$$

The 2-step ahead forecast is

$$\sigma_t^2(2) = \alpha_0 + \alpha_1 \sigma_t^2(1) + \alpha_2 r_t^{*2} + ... + \alpha_p r_{t+2-p}^{*2}.$$

S-Plus Demonstration

Output edited and % marks explanation.

```
> arch3.fit=garch(intc~1,~garch(3,0))
> summary(arch3.fit)
Call:
garch(formula.mean = intc ~ 1, formula.var = ~ garch(3, 0))

Mean Equation: intc ~ 1
Conditional Variance Equation:  ~ garch(3, 0)
Conditional Distribution:  gaussian
-------------------------------------------------------------
Estimated Coefficients:
-------------------------------------------------------------
           Value Std.Error t value Pr(>|t|)
       C 0.01713  0.006626  2.5860 0.005047  % one-sided
       A 0.01199  0.001107 10.8325 0.000000  % p-value
 ARCH(1) 0.17874  0.080294  2.2260 0.013309
 ARCH(2) 0.07720  0.050552  1.5271 0.063800
 ARCH(3) 0.05722  0.076928  0.7438 0.228747
-------------------------------------------------------------
> arch1=garch(intc~1,~garch(1,0)) % A simplified model
> summary(arch1)
Call:
garch(formula.mean = intc ~ 1, formula.var = ~ garch(1,0))

Mean Equation: intc ~ 1
Conditional Variance Equation:  ~ garch(1, 0)
Conditional Distribution:  gaussian
-------------------------------------------------------------
Estimated Coefficients:
-------------------------------------------------------------
           Value Std.Error t value  Pr(>|t|)
       C 0.01741  0.006231   2.794 2.737e-03
       A 0.01258  0.001246  10.091 0.000e+00
 ARCH(1) 0.35258  0.088515   3.983 4.094e-05
-------------------------------------------------------------
> stdresi=arch1$residuals/arch1$sigma.t   % Standardized
> autocorTest(stdresi,lag=10)             % residuals


Null Hypothesis: no autocorrelation
Test Statistics:
Test Stat 13.7820  p.value  0.1832

Dist. under Null: chi-square with 10 degrees of freedom
> archTest(stdresi,lag=10)  % ARCH test for residuals

Null Hypothesis: no ARCH effects
Test Statistics:
Test Stat 11.3793  p.value  0.3287

Dist. under Null: chi-square with 10 degrees of freedom
> arch1$asymp.sd  % Obtain unconditional variance
[1] 0.1393796

> plot(arch1)  % Obtain various plots, including the
               % fitted volatility series.
```

5.5 Examples of ARCH modeling

Example 3.1. We shall build an ARCH model for the monthly log returns of Intel stock. The sample ACF and PACF in Figure 3.1 show the presence of conditional heteroscedasticity. It is confirmed by test for ARCH effects.

We proceed to identify the order m of ARCH model. The PACF in Figure 3.1 (d) indicates that an ARCH(3) model might be appropriate. No correlation in r_t. So we specify the model:

$$r_t = \mu + r_t^*, \quad r_t^* = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^{*2} + \alpha_2 r_{t-2}^{*2} + \alpha_3 r_{t-3}^{*2}$$

for monthly returns of Intel stock. Assuming that ε_t are iid standard normal, we obtain the fitted model:

$$r_t = 0.0171 + r_t^*, \qquad \sigma_t^2 = 0.0120 + 0.178 r_{t-1}^{*2} + 0.0772 r_{t-2}^{*2} + 0.0572 r_{t-3}^{*2},$$

where standard errors for parameters are $0.0066, 0.0011, 0.0803, 0.0506$ and 0.0769, see output below. The estimates for α_2 and α_3 are statistically non-significant at 5-percent level. Therefore the model can be simplified to

$$r_t = 0.0174 + r_t^*, \qquad \sigma_t^2 = 0.0126 + 0.352 r_{t-1}^{*2},$$

where the standard errors are $0.0062, 0.0012$ and 0.0885, respectively.
All estimates are highly significant.
To check goodness of fit we have to investigate residuals

$$\tilde{\varepsilon}_t = \frac{r_t^*}{\sigma_t}.$$

Figure 3.4 shows the residuals and sample ACF of some functions of $\tilde{\varepsilon}_t$.

To check that $\tilde{\varepsilon}_t$ is an iid sequence Ljung-Box statistics $Q(10)$ gives p-values 0.18, and for $\tilde{\varepsilon}_t^2$, statistics $Q(10) = 11.38$ has p-value 0.33, see output. Consequently, ARCH(1) model is an adequate model for describing the data at 5-percent significance level.

Equation above shows that the Intel stock has 1.74 -percent monthly return which is very high. Secondly, $\hat{\alpha}_1^2 = 0.353^2 < 1/3$, so that the ARCH(1) model is stationary. and finite forth moment exists. The unconditional standard deviation is

$$\frac{\alpha_0}{1-\alpha_1} = \sqrt{0.0126(1 - 0.352)} = 0.1394.$$

ARCH(1) model can be used for prediction of volatility σ_t^2 of Intel stock.

Example 3.2 Figure 3.2a shows the percentage exchange rate between mark and dollar in 10-minute intervals, we discussed before. Series does not have serial correlation. Sample PACF indicates big spikes at lag 1 and 3. Similarly as in Example 3.1 we specify model ARCH(3) for this data. Unconditional Gaussian maximum likelihood estimation gives the following model:

$$r_t = 0.0018 + r_t^*$$

and

$$r_t^* = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = 0.0022 + 0.322 r_{t-1}^{*\,2} + 0.074 r_{t-2}^{*\,2} + 0.093 r_{t-3}^{*\,2}.$$

All estimates are significant at 5% level. Model check for residuals $\tilde{\varepsilon}_t$ indicates that the estimated ARCH(3) model is adequate.

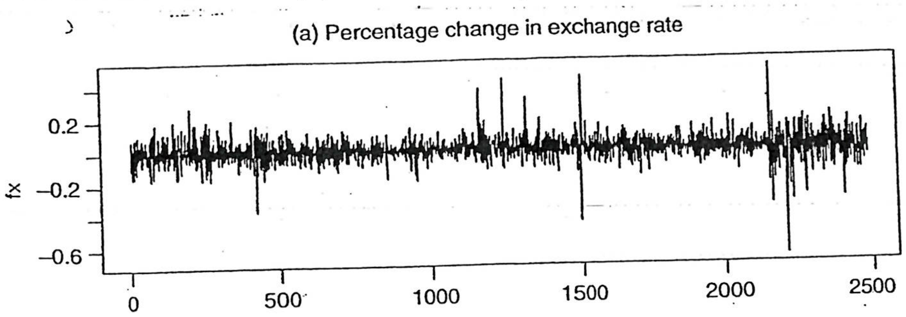


Figure 3.2. (a) Time plot of 10-minute returns of the exchange rate between Deutsche mark and U.S. dollar and (b) the squared returns.

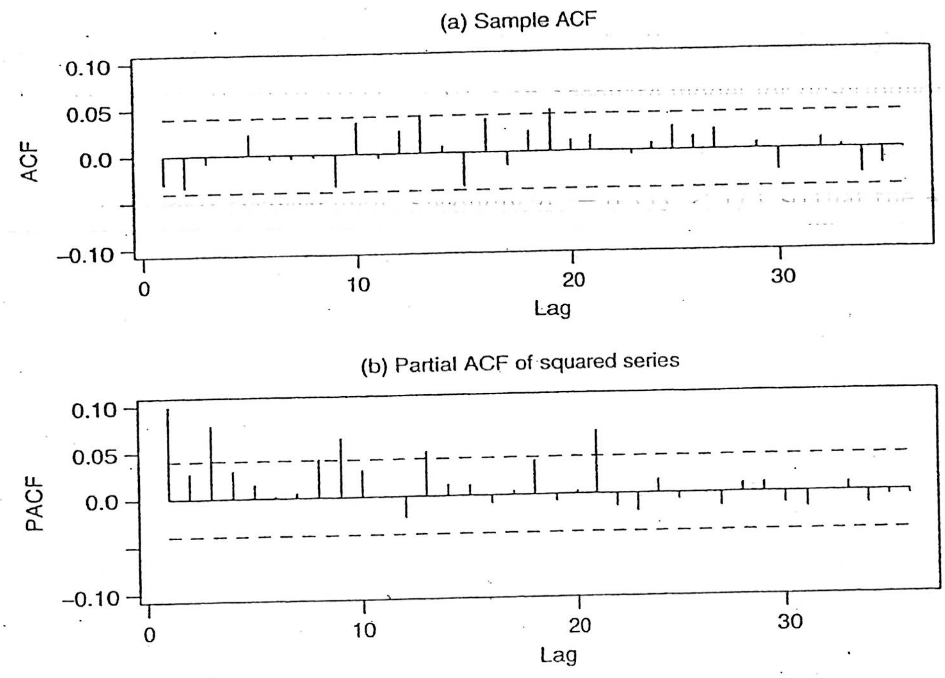


Figure 3.3. (a) Sample autocorrelation function of the return series of mark/dollar exchange rate and (b) sample partial autocorrelation function of the squared returns.

5.6 The GARCH model

The drawback of ARCH(p) model is that often in applications, to describe adequately model, it requires many parameters. We shall have later an example, where the model ARCH(9) is needed for modeling stochastic volatility.

Let r_t be a log returns series, and

$$r_t = \mu_t + r_t^*,$$

where

$$\mu_t = E[r_t | F_{t-1}], \quad r_t^* = r_t - \mu_t$$

is the conditional mean.

GARCH(p,q) model

We say that the innovations a_t follow GARCH(p,q) model if

$$r_t^* = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^{*2} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{20}$$

where ε_t are i.i.d. random variables with mean 0 and variance 1,

$$\alpha_0 \geq 0, \alpha_1 \geq 0, ..., \alpha_p \geq 0, \quad \beta_1 \geq 0, ..., \beta_q \geq 0 \text{ are parameters.}$$

Stationary solution. To assure existence of a stationary solution with a finite variance we impose condition

$$\alpha_1 + ... + \alpha_p + \beta_1 + ... + \beta_q < 1. \tag{21}$$

In applications often it is assumed that errors ε_t have standard normal or standardized Student-t distribution.

When $q = 0$, the GARCH(p,0) model reduces to ARCH(p) model.

ARMA type representation of GARCH model.

We set

$$\eta_t = r_t^{*2} - \sigma_t^2.$$

Then η_t are uncorrelated but dependent variables. Then, using $\sigma_t^2 = r_t^{*2} - \eta_t$, we obtain,

$$\begin{aligned} r_t^{*2} &= \sigma_t^2 + \eta_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^{*2} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \eta_t \\ &= \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^{*2} + \sum_{j=1}^{q} \beta_j r_{t-j}^{*2} + \eta_t - \sum_{j=1}^{q} \beta_j \eta_{t-j}. \end{aligned}$$

Equation is an ARMA form for squares r_t^{*2} where η_t plays role of a white noise. Using formula for unconditional mean of an ARMA model, we have

$$E[r_t^{*2}] = \frac{\alpha_0}{1 - \alpha_1 - \ldots - \alpha_p - \beta_1 - \ldots - \beta_q}$$

where condition (21) implies that the denominator is positive and therefore variance $Var(r_t^{*2})$ is finite.

GARCH(1,1) model

The simple GARCH model is GARCH (1,1) model where condition variance is modeled by

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^{*2} + \beta_1 \sigma_{t-1}^2, \quad 0 \le \alpha_1, \beta_1 \le 1, \quad \alpha_1 + \beta_1 < 1.$$

Properties:

- A large r_{t-1}^{*2} or σ_{t-1}^2 leads to a large σ_t^2. This means that large r_{t-1}^{*2} is followed by another large r_t^{*2}, so we have clustering property, common in financial time series.

- The excess kurtosis is greater then 3:
  under condition $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, we have

  $$\frac{E[r_t^{*4}]}{(E[r_t^{*2}])^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3.$$

  This shows that the tail of GARCH(1,1) distribution is heavier than that of a normal distribution.

- GARCH(1,1) model is a simple parametrical model which can be used to describe volatility evolution.

Forecasting of GARCH(1,1) model. Forecasting of GARCH models is similar to that of ARMA models. To obtain 1-step ahead forecast of volatility σ_t^2, note that

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 r_t^{*2} + \beta_1 \sigma_t^2,$$

where r_t^{*2}, σ_t^2 are know at time t and parameters $\alpha_0, \alpha_1, \beta_1$ are estimated from the data.

1-step ahead forecast at horizon t is

$$\sigma_t^2(1) = E[\sigma_{t+1}^2 | F_t] = \alpha_0 + \alpha_1 r_t^{*2} + \beta_1 \sigma_t^2.$$

2-step ahead forecast. We rewrite the volatility equation as

$$\sigma_{t+2}^2 = \alpha_0 + (\alpha_1 + \beta_1) r_{t+1}^{*2} + \alpha_1 \sigma_{t+1}^2 (\varepsilon_{t+1}^2 - 1).$$

Since $E(\varepsilon_{t+1}^2 - 1 | F_t) = 0$ then

$$\sigma_t^2(2) = E[\sigma_{t+2}^2 | F_t] = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2(1).$$

k-step ahead prediction: for $k \geq 2$, is

$$\sigma_t^2(k) = E[\sigma_{t+k}^2 | F_t] = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2(k-1).$$

By repeated substitutions in the above equation we obtain the formula

$$\sigma_t^2(k) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{k-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{k-1} \sigma_t^2(1).$$

We see that

$$\sigma_t^2(k) \to \frac{\alpha_0}{1 - \alpha_1 - \beta_1} = E[\sigma_t^2] = E[r_t^{*2}]$$

if $\alpha_1 + \beta_1 < 1$, and therefore GARCH(1,1) process is mean reverting.

The literature on GARCH models is very rich.

The weakness of GARCH model is the same as in case of ARCH model:

- The model responses equally to positive and negative shocks
- The tails of distribution of GARCH models are not sufficiently heavy comparing to some financial data.

5.7 Illustrative example

The modeling procedure used in ARCH case can be used also for GARCH models. However, specifying the order of GARCH models is not an easy task.

In practise, only the lower order models are used:

GARCH(1,1), GARCH(1,2), GARCH(2,1).

Parameters can be estimated using conditional maximum likelihood method assuming that that initial value σ_1^2 of volatility is fixed (known). Then, in GARCH(1,1) case, the volatility σ_t^2 can be computed recursively.

- In some applications a good starting value of σ_1^2 might be the sample variance of r_t^{*2}.
- The fitted model, its goodness, can be checked by using standardized residuals

$$\tilde{\varepsilon}_t = \frac{r_t^*}{\sigma_t}$$

and its squared model

$$\tilde{\varepsilon}_t^2 = \frac{r_t^{*2}}{\sigma_t^2}.$$

If case of good fit, there should be no ARCH effect in residuals $\tilde{\varepsilon}_t$, i.e.

– $\tilde{\varepsilon}_t$ should be uncorrelated,

– $\tilde{\varepsilon}_t^2$ should be uncorrelated.

Example We shall discuss the monthly excess returns of and S&P index from 1926 for 792 observations. Denote excess return r_t

Definition: Excess return of an asset at time t is the difference between the assets return and some reference asset:

$$r_t = r_{t,asset} - r_{t,ref.asset}.$$

The reference asset is often taken to be a riskless such as a short-term US treasury bill returns.

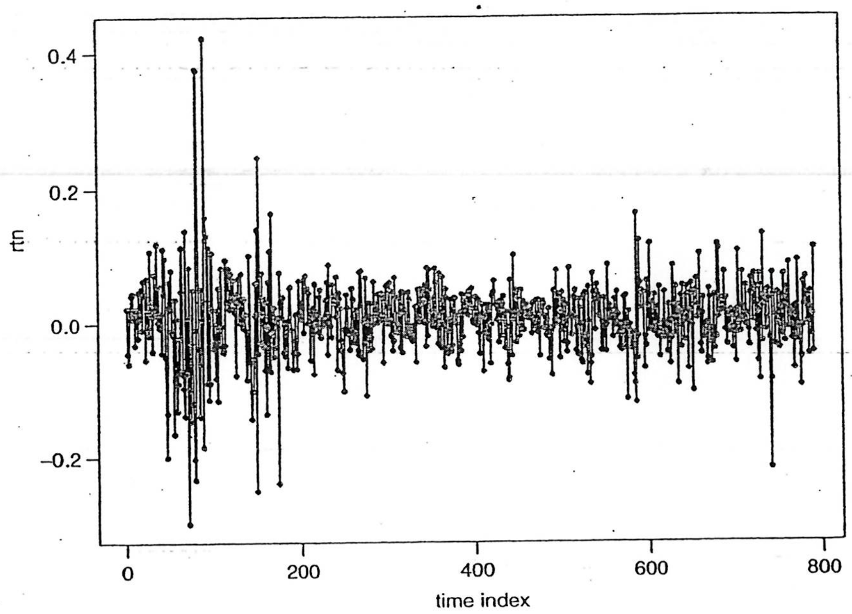


Figure 3.5. Time series plot of the monthly excess returns of the S&P 500 index.

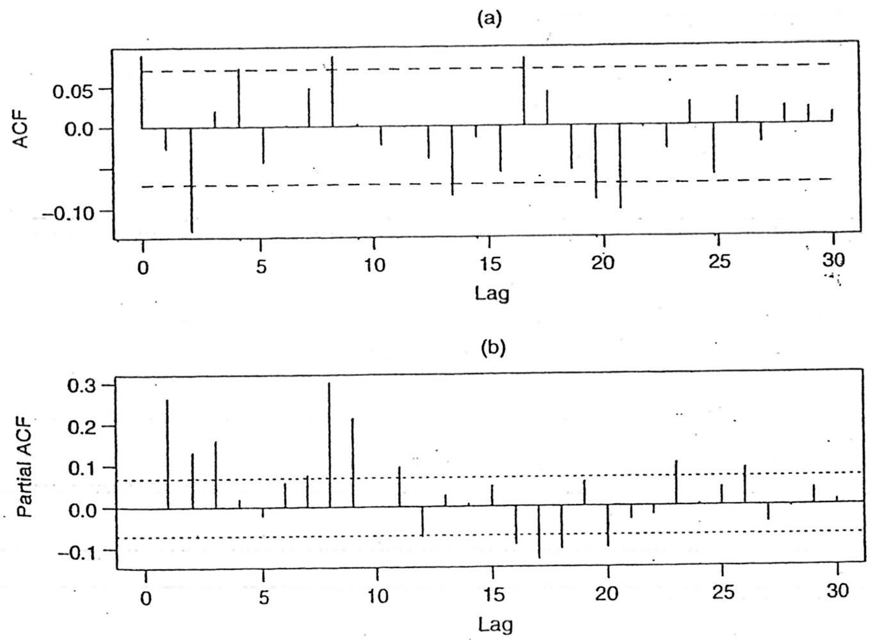


Figure 3.6. (a) Sample ACF of the monthly excess returns of the S&P 500 index and (b) sample PACF of the squared monthly excess returns.

In Figure 3.6 we see the sample ACF of r_t, and the sample PACF of r_t^2:

- the ACF has serial correlation at lags 1 and 3;
- the PACF shows strong linear dependence (correlation) for a number of lags. That means that for r_t^2 ARCH models do not fit well, and we shall fit GARCH(1,1) model.

Model The ACF suggests to fit for r_t MA(3) model. The estimated model:

$$r_t = 0.0062 + r_t^* + 0.0944 r_{t-1}^* - 0.1407 r_{t-3}^* + e_t, \qquad \hat{\sigma}_{r^*}^2 = 0.0576.$$

where all coefficients are significant at 5 % level, and e_t in case of 'good' fit should be a white noise.

To simplify dependence on r_t^*, instead of MA(3) model we shall fit to the data an AR(3) model

$$r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + \phi_0 + r_t^*$$

where $\phi_1, \phi_2, \phi_3, \phi_0$ are unknown parameters and ar_t^* is a white noise. The fitted model, assuming normality of r_t^* is

$$r_t = 0,088 r_{t-1} - 0.023 r_{t-2} - 0.123 r_{t-3} + 0.0066 + r_t^*, \quad \hat{\sigma}_{r^*}^2 = 0.00333.$$

Although the residuals r_t^* might be a white noise (uncorrelated), they still might be dependent. To model this type of data (so called GARCH effect), we use GARCH(1,1) model for σ_t:

$$r_t^* = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 r_{t-1}^{*2}.$$

A joint estimation of the AR(3)-GARCH(1,1) model gives

$$r_1 = 0.0078 + 0.032 r_{t-1} - 0.029 r_{t-2} - 0.008 r_{t-3} + r_t^*,$$

$$\sigma_t^2 = 0.000084 + 0.121 r_{t-1}^{*2} + 0.8523 \sigma_{t-1}^2.$$

Having estimates of parameters we can estimate the unconditional variance of r_t^*:

$$Var(r_t^*) \sim \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)} = \frac{0.000084}{1 - 0.8523 - 0.1213} = 0.00317$$

which is close to that derived fitting AR(3) model to r_t.

t-ratios suggest that all three parameter in the mean (AR(3) part) are not significant. We refine the model, dropping those parameters. The estimated model then is

$$r_t = 0.0076 + r_t^*, \qquad \sigma_t^2 = 0.000086 + 0.1216 r_{t-1}^{*2} + 0.8511 \sigma_{t-1}^2.$$

The standard error of the constant in the mean equation is 0.0015, so parameter 0.0076 is significant at 5% level.

The standard errors of the parameters in the volatility equation are 0.000024, 0.0197, and 0.0190, respectively. These parameter are significant. The unconditional variance of r_t^* is

$$\frac{0.000086}{1 - 0.8511 - 0.1216} = 0.00314.$$

This is a simple GARCH(1,1) model.

Goodness-of fit To check the goodness of fit of the estimated model we compute standardized shocks

$$\tilde{\varepsilon}_t = \frac{r_t^*}{\sigma_t}.$$

– Figure 3.7 shows the estimated volatility σ_t^2 and the standardized shocks $\tilde{\varepsilon}_t$.

– Figure 3.8 shows the sample ACF of the residuals $\tilde{\varepsilon}_t$ and squares $\tilde{\varepsilon}_t^2$. The ACF does not show any correlation in $\tilde{\varepsilon}_t$ and $\tilde{\varepsilon}_t^2$, which indicates a good fit of the model.

The Ljung-Box statistics takes value

– Q(12)=11.99 with p-value 0.45 for $\tilde{\varepsilon}_t$, and

– $Q(12) = 11.11$ with p-value 0.36 for $\tilde{\varepsilon}_t^2$.

The model appears to be adequate, describing the dependence structure in the return and volatility series.

Note: The fitted model has property

$$\hat{\alpha}_1 + \hat{\beta}_1 = 0.9772$$

is close to 1. This phenomenon is often observed in practice and leads to the model GARCH(1,1) with constrain $\alpha_1 + \beta_1 = 1$. The GARCH(1,1) model with this property is called IGARCH model.

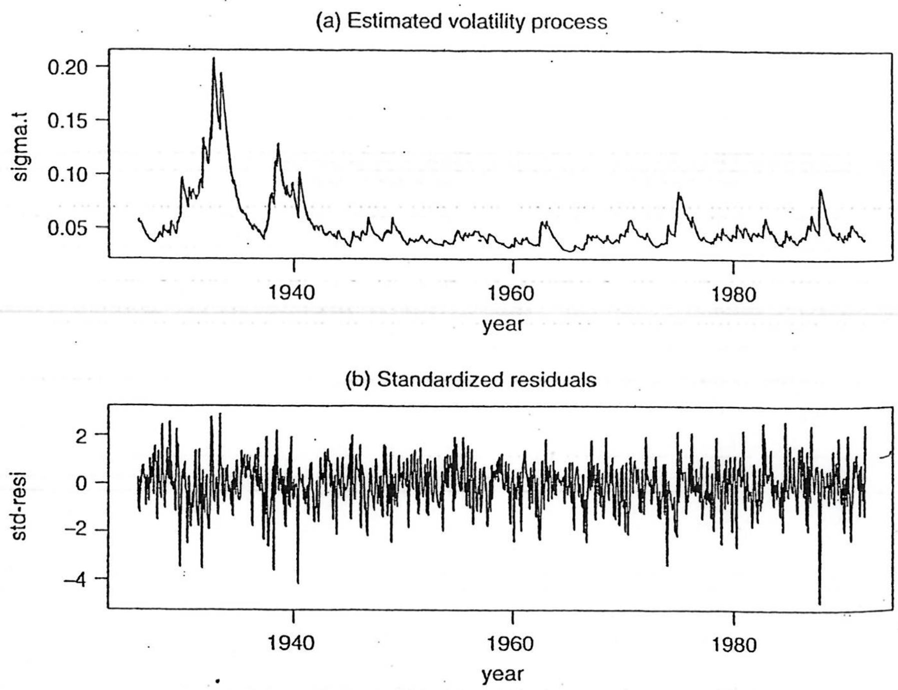


Figure 3.7. (a) Time series plot of estimated volatility (σ_t) for the monthly excess returns of the S&P 500 index and (b) the standardized shocks of the monthly excess returns of the S&P 500 index. Both plots are based on the GARCH(1,1) model in Eq. (3.18).

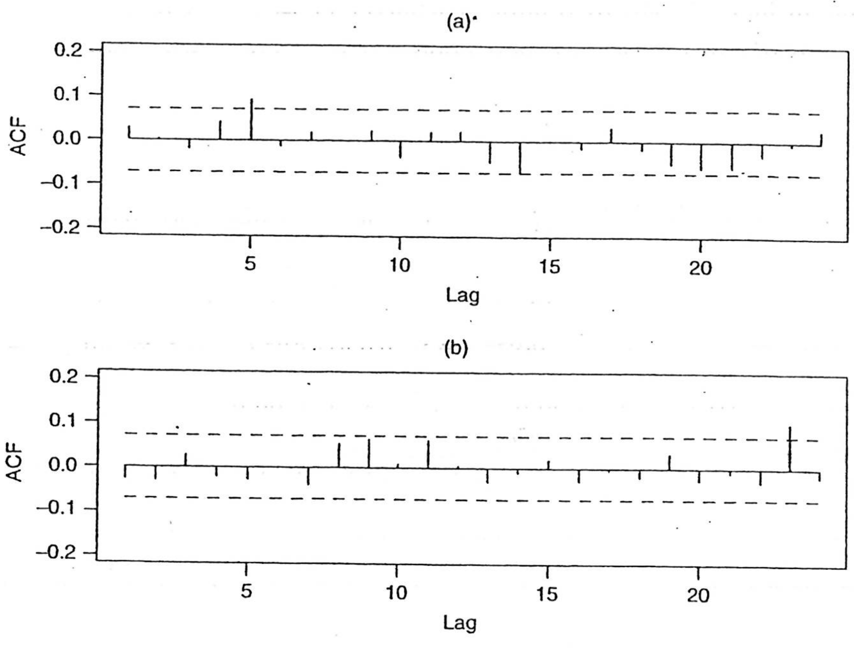


Figure 3.8. Model checking of the GARCH(1,1) model in Eq. (3.18) for monthly excess returns of the S&P 500 index: (a) sample ACF of standardized residuals and (b) sample ACF of the squared standardized residuals.

Forecasting: To forecast the volatility of monthly excess returns of the S&P index, we use the volatility equation

$$r_1 = 0.0076 + r_t^*, \qquad \sigma_t^2 = 0.000086 + 0.1216 r_{t-1}^{*2} + 0.8511 \sigma_{t-1}^2.$$

Then the 1-set ahead forecast is

$$\sigma_t^2(1) = 0.00008 + 0.121 r_t^{*2} + 0.8511 \sigma_t^2$$

where r_t^{*2} is is the residual, in our case $r_t^* = r_t - 0.0076$, and σ_t^2 can be computed using GARCH(1,1) equation for volatility. Computing it, we can fix σ_0^2 at zero or at unconditional variance of r_t^*. For multi-step forecast, we use recursive formula.

Table 3.2 shows some mean and volatility forecasts for the returns of the S&P index with forecast horizon $t = 792$.

Table 3.1. Volatility Forecasts for the Monthly Excess Returns of the S&P 500 Index[a]

Horizon	1	2	3	4	5	∞
Return	0.0076	0.0076	0.0076	0.0076	0.0076	0.0076
Volatility	0.0536	0.0537	0.0537	0.0538	0.0538	0.0560

[a]The forecast origin is $h = 792$, which corresponds to December 1991. Here volatility denotes conditional standard deviation.

excess return of the S&P 500 index with forecast origin $h = 792$ based on the GARCH(1,1) model in Eq. (3.18).

Some S-Plus Commands Used in Example 3.3

```
> fit=garch(sp~ar(3),~garch(1,1))
> summary(fit)
> fit=garch(sp~1,~garch(1,1))
> summary(fit)
> names(fit)
 [1] "residuals" "sigma.t"     "df.residual" "coef" "model"
 [6] "cond.dist" "likelihood" "opt.index"    "cov"
[10] "prediction" "call"      "asymp.sd"    "series"
> % Next, compute the standardized residuals
> stdresi=fit$residuals/fit$sigma.t
> autocorTest(stdresi,lag=24)
> autocorTest(stdresi^2,lag=24)
> predict(fit,5) % Compute predictions
```

Note that in the prior commands the volatility series σ_t is stored in `fit$sigma.t` and the residual series of the returns in `fit$residuals`.

S-Plus Commands Used

```
> fit1 = garch(sp~1,~garch(1,1),cond.dist='t',cond.par=5,
+ cond.est=F)
> summary(fit1)
> stresi=fit1$residuals/fit1$sigma.t
> autocorTest(stresi,lag=10)
> autocorTest(stresi^2,lag=10)
```

t Innovations Assume now that in GARCH(1,1) equation

$$r_t^* = \sigma_t \varepsilon_t,$$

the errors ε_t follow a Standardized Student-t distribution with 5 degrees of freedom. Re-estimating GARCH(1,1) model, we obtain

$$r_1 = 0.0085 + r_t^*, \qquad \sigma_t^2 = 0.00012 + 0.11216 r_{t-1}^{*2} + 0.8432 \sigma_{t-1}^2,$$

where standard errors of parameters are 0.0015, 0.51×10^{-4}, 0.0296 and 0.0371. This model is essentially IGARCH model as $\hat{\alpha}_1 + \hat{\beta}_1 = 0.95$ is close to 1.

Goodness of fit: The Ljung-Box statistics of the residuals gives:
— $Q(10) = 11.38$ with p-value 0.33 for $\tilde{\varepsilon}_t$; and
— $Q(10) = 10.48$ with p-value 0.40 for $\tilde{\varepsilon}_t^2$.
Thus this GARCH(1,1) model is adequate as well as the model where we used normally distributed errors ε_t.