Ch 6 - part 2

- In part 1, we discussed the ARCH(1) model:

$$
\begin{aligned}
r_t &= \sigma_{t|t-1}\varepsilon_t, \\
\sigma^2_{t|t-1} &= \omega + \alpha r^2_{t-1}.
\end{aligned}
$$

- We discussed properties of the (unconditional) distribution of return r_t for $\omega > 0$ and $0 \leq \alpha < 1$.
- Next: parameter estimation and forecasting.

Estimating ARCH(1) model parameters and conditional variance

$$
\begin{aligned}
r_t &= \sigma_{t|t-1}\varepsilon_t, \\
\sigma^2_{t|t-1} &= \omega + \alpha r^2_{t-1}.
\end{aligned}
$$

- Maximum likelihood estimation is used to estimate the model parameters in the ARCH(1) model.
  - If $\varepsilon_t \sim N(0,1)$, and if we condition on r_1, the likelihood function for the r_t's can be obtained based on (switching to previous notation):

$$R_t | R_{t-1} = r_{t-1} \sim N(0, \omega + \alpha r^2_{t-1}).$$

- After estimating the model parameters, estimates of the conditional variance are given by:

$$\hat{\sigma}^2_{t|t-1} = \hat{\omega} + \hat{\alpha} r^2_{t-1},$$

where $\hat{\sigma}_{1|0}$ is set equal to r_1^2 or to the unconditional variance of r_t.

Simulation example ARCH(1) continued

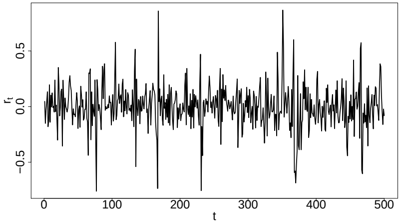
- As before, ARCH(1) model:

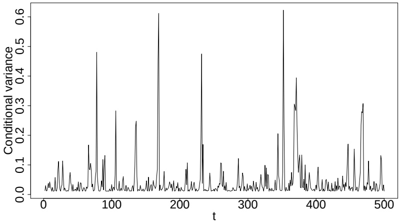
$$
\begin{aligned}
r_t &= \sigma_{t|t-1}\varepsilon_t, \\
\sigma^2_{t|t-1} &= \omega + \alpha r^2_{t-1},
\end{aligned}
$$

  where $\varepsilon_t \sim N(0,1)$, $\omega = 0.01$ and $\alpha = 0.9$.

- Estimates of the conditional variance are given by:

$$
\hat{\sigma}_{t|t-1} = \hat{\omega} + \hat{\alpha} r^2_{t-1}.
$$





How to forecast conditional variance?

- Let's go back to notation for random variables versus observed values:
  Define the forecast for conditional variance as
  $\hat{v}_{t+g|t} = Var(R_{t+g}|R_j = r_j, j = 1, 2, \dots t)$.
- $Var(R_{t+g}|R_j = r_j, j = 1, 2, \dots t) = E(R^2_{t+g}|R_j = r_j, j = 1, 2, \dots t)$ because $E(R_{t+g}|R_j = r_j, j = 1, 2, \dots t) = 0$.
- It follows that

$$
\begin{aligned}
\hat{v}_{t+g|t} &= E(R^2_{t+g}|R_j = r_j, j = 1, 2, \dots t), \\
&= E(V_{t+g|t+g-1}\varepsilon^2_{t+g}|R_j = r_j, \text{ for } j = 1, \dots, t), \\
&= E(\varepsilon^2_{t+g})E(V_{t+g|t+g-1}|R_j = r_j, \text{ for } j = 1, \dots, t), \\
&= E(V_{t+g|t+g-1}|R_j = r_j, \text{ for } j = 1, \dots, t), \\
&= E(\omega + \alpha R^2_{t+g-1}|R_j = r_j, \text{ for } j = 1, \dots, t), \\
&= \omega + \alpha E(R^2_{t+g-1}|R_j = r_j, \text{ for } j = 1, \dots, t).
\end{aligned}
$$

How to forecast conditional variance?

- We found:

$$\hat{v}_{t+g|t} = \omega + \alpha E(R^2_{t+g-1} | R_j = r_j, \text{ for } j = 1, \ldots, t).$$

- For $g = 1$:

$$\begin{aligned} \hat{v}_{t+1|t} &= E(\omega + \alpha R^2_t | R_j = r_j, \text{ for } j = 1, \ldots, t), \\ &= \omega + \alpha r^2_t. \end{aligned}$$

  - Note that $\omega = \sigma^2(1 - \alpha)$ (where σ^2 was the unconditional variance for r_t), so the forecasted conditional variance is a weighted average of the long-run variance and the current squared return.

- For $g > 1$:

$$\begin{aligned} \hat{v}_{t+g|t} &= \omega + \alpha E(R^2_{t+g-1} | R_j = r_j, \text{ for } j = 1, \ldots, t), \\ &= \omega + \alpha \hat{v}_{t+g-1|t}. \end{aligned}$$

- Forecast intervals for r_{t+g} are reported as $\hat{r}_{t+g} \pm 1.96 \sigma_{t+g|t}$, where $\hat{r}_{t+g} = E(R_{t+g} | R_j = r_j, \text{ for } j = 1, \ldots, t) = 0$.

Fancier models: Inclusion of more predictors

- Volatility may depend on additional terms, e.g., squared returns at time $t-2, t-3, \ldots$.
- Model extensions (still assuming $r_t = \sigma_{t|t-1}\varepsilon_t$):
  - ARCH(q)

$$\sigma^2_{t|t-1} = \omega + \alpha_1 r^2_{t-1} + \alpha_2 r^2_{t-2} + \ldots + \alpha_q r^2_{t-q}$$

  - Generalized autoregressive conditional heteroskedasticity model, GARCH(p,q) (note: sometimes p and q are swapped in software)

$$\begin{aligned}\sigma^2_{t|t-1} &= \omega + \beta_1 \sigma^2_{t-1|t-2} + \ldots + \beta_p \sigma^2_{t-p|t-p-1} \\ &\quad + \alpha_1 r^2_{t-1} + \alpha_2 r^2_{t-2} + \ldots + \alpha_q r^2_{t-q}\end{aligned}$$

- Parameter restrictions are needed to guarantee positive variance and stationarity (but not all parameters need to be positive to guarantee a positive variance).

Investigating the GARCH(p, q) process

- Again let $\eta_t = r_t^2 - \sigma^2_{t|t-1}$, then plug in $\sigma^2_{t|t-1} = r_t^2 - \eta_t$ and find:

$$\begin{aligned}
\sigma^2_{t|t-1} &= \omega + \beta_1 \sigma^2_{t-1|t-2} + \ldots + \beta_p \sigma^2_{t-p|t-p-1} \\
&\quad + \alpha_1 r^2_{t-1} + \alpha_2 r^2_{t-2} + \ldots + \alpha_q r^2_{t-q}, \\
r_t^2 &= \omega + (\beta_1 + \alpha_1) r^2_{t-1} \\
&\quad + \ldots + (\beta_{max(p,q)} + \alpha_{max(p,q)}) r^2_{t-max(p,q)} \\
&\quad + \eta_t - \beta_1 \eta_{t-1} - \ldots - \beta_p \eta_{t-p},
\end{aligned}$$

  where $\beta_k = 0$ for $k > p$ and $\alpha_k = 0$ for $k > q$.
- As before, the η_t's have mean zero and not autocorrelated and not correlated with the squared returns:
  - A GARCH(p, q) model for r_t corresponds to an ARMA($\max(p, q), p$) model for r_t^2.

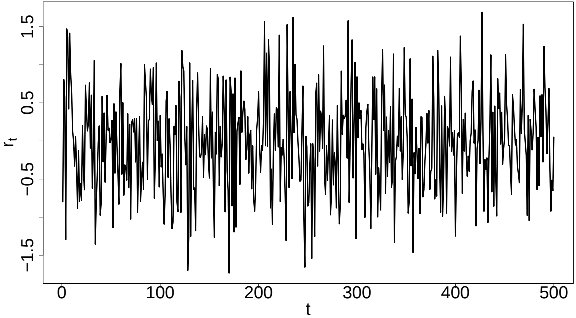
GARCH(1,1) simulation example

- GARCH(1,1) model:

$$\begin{aligned} r_t &= \sigma_{t|t-1}\varepsilon_t, \\ \sigma^2_{t|t-1} &= \omega + \alpha r^2_{t-1} + \beta\sigma^2_{t-1|t-2}, \end{aligned}$$

  where $\varepsilon_t \sim N(0,1)$, $\omega = 0.02$, $\alpha = 0.05$ and $\beta = 0.9$.
- R-code:

```
garch11.sim=garch.sim(alpha=c(.02,.05),beta = 0.9, n=500)
```



Estimating GARCH(p, q) model parameters and conditional variance

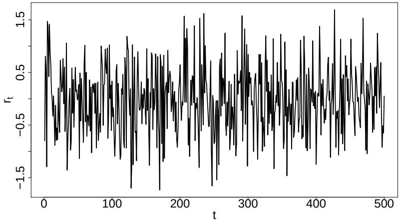
- As for ARCH(1), we can use maximum likelihood estimation to estimate the model parameters in the GARCH(p, q) model.
- For the simulated data with $\omega = 0.02$, $\alpha = 0.05$ and $\beta = 0.9$, I got convergence issues when using the code from the book (using "garch"), so I recommend using garchFit from the fGarch package.

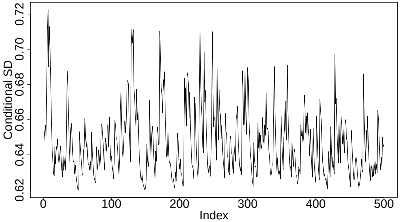
```
m11=garch(x=garch11.sim,order=c(1,1))
...
 ***** FALSE CONVERGENCE *****
...
    Estimate  Std. Error  t value Pr(>|t|)
a0 3.961e-01   4.081e-01    0.971    0.332
a1 5.601e-02   6.275e-02    0.893    0.372
b1 3.610e-13   9.857e-01    0.000    1.000
```

Estimating GARCH(p, q) model parameters and conditional variance

- For the simulated data with $\omega = 0.02$, $\alpha = 0.05$ and $\beta = 0.9$:

```
library(fGarch)
m11=garchFit(x=garch11.sim,order=c(1,1))
        Estimate  Std. Error  t value Pr(>|t|)
omega     0.13880      0.29796     0.466     0.641
alpha1    0.03351      0.04934     0.679     0.497
beta1     0.63481      0.74102     0.857     0.392
```





Forecasting for GARCH$(1, 1)$

- Approach is the same as for the ARCH(1) model:

$$
\begin{aligned}
\hat{v}_{t+g|t} &= E(R^2_{t+g}|R_j = r_j, j = 1, 2, \ldots t), \\
&= E(V_{t+g|t+g-1}\varepsilon^2_{t+g}|R_j = r_j,\ \text{for } j = 1, \ldots, t), \\
&= E(\varepsilon^2_{t+g})E(V_{t+g|t+g-1}|R_j = r_j,\ \text{for } j = 1, \ldots, t), \\
&= E(V_{t+g|t+g-1}|R_j = r_j,\ \text{for } j = 1, \ldots, t), \\
&= E(\omega + \alpha_1 R^2_{t+g-1} + \beta_1 V_{t+g-1|t+g-2}|R_j = r_j,\ \text{for } j = ...), \\
&= \omega + \alpha_1 E(R^2_{t+g-1}|R_j = r_j,\ \text{for } j = 1, \ldots, t) \\
&\quad + \beta_1 E(V_{t+g-1|t+g-2}|R_j = r_j,\ \text{for } j = 1, \ldots, t).
\end{aligned}
$$

Forecasting for GARCH$(1,1)$

- We found

$$\begin{aligned}\hat{v}_{t+g|t} &= \omega + \alpha_1 E(R^2_{t+g-1}|R_j = r_j,\ \text{for } j = 1, \ldots, t) \\ &\quad + \beta_1 E(V_{t+g-1|t+g-2}|R_j = r_j,\ \text{for } j = 1, \ldots, t).\end{aligned}$$

- For $g = 1$:

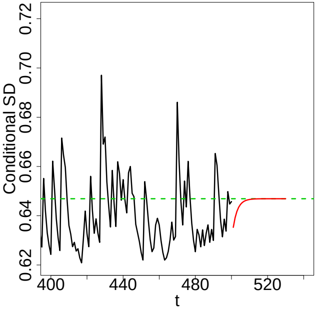
$$\begin{aligned}\hat{v}_{t+1|t} &= \omega + \alpha_1 E(R^2_t|R_j = r_j,\ \text{for } j = 1, \ldots, t) \\ &\quad + \beta_1 E(V_{t|t-1}|R_j = r_j, \ldots), \\ &= \omega + \alpha_1 r_t^2 + \beta_1 v_{t|t-1},\end{aligned}$$

  where we assume $v^2_{t|t-1}$ is known (estimated).

- For $g > 1$:

$$\begin{aligned}\hat{v}_{t+g|t} &= \omega + \alpha_1 E(R^2_{t+g-1}|R_j = r_j,\ \text{for } j = 1, \ldots, t) \\ &\quad + \beta_1 E(V_{t+g-1|t+g-2}|R_j = r_j,\ \text{for } j = 1, \ldots, t), \\ &= \omega + \alpha_1 \hat{v}_{t+g-1|t} + \beta_1 \hat{v}_{t+g-1|t}, \\ &= \omega + (\alpha_1 + \beta_1)\hat{v}_{t+g-1|t}.\end{aligned}$$

Example of GARCH(1,1) forecast: SD



Relevant info for $\sigma_{n+1|n}$:

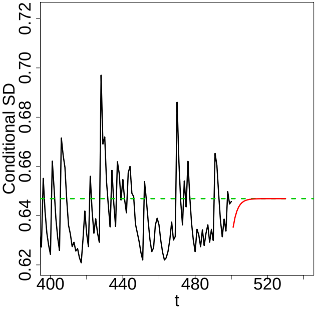
```
> coef(m11)
     omega     alpha1      beta1
0.13880459 0.03350646 0.63481171
> m11@sigma.t[n]
[1] 0.6457404
> garch11.sim[n]
[1] 0.05024388
```

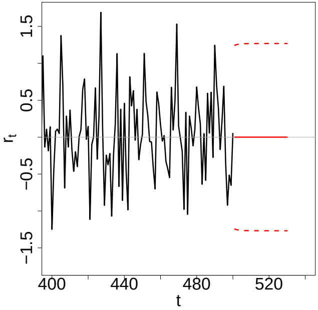
- The conditional SD converges to the unconditional SD σ for r_t: here $\hat{\sigma} = \sqrt{\hat{\omega}/(1 - \hat{\alpha} - \hat{\beta})} \approx 0.65$ (green dotted line).
  - This follows from the ARMA(1,1) representation for r_t:

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

thus $\sigma^2 = E(r_k^2) = \omega/(1 - \beta_1 - \alpha_1)$.

Example of GARCH(1,1) forecast: Returns





- If $r_t = \sigma_{t|t-1}\varepsilon_t$ with $\varepsilon_t \sim N(0,1)$, the forecast intervals for r_{t+h} are given by $\hat{r}_{t+h} \pm 1.96\sigma_{t+h|t}$, where $\hat{r}_{t+h} = E(r_{t+h}|r_1, \ldots, r_t) = 0$.

Summary so far

- Our goal is to model and forecast asset volatility $\sigma_{t|t-1} = SD(r_t|r_{t-1}, r_{t-2}, \ldots)$, the conditional SD for returns r_t for a financial asset.
- To capture changes in volatility, we discussed the class of GARCH(p, q) models, e.g. the GARCH(1,1) model is given by

$$\begin{aligned} r_t &= \sigma_{t|t-1}\varepsilon_t, \\ \sigma^2_{t|t-1} &= \omega + \alpha r^2_{t-1} + \beta\sigma^2_{t-1|t-2}, \end{aligned}$$

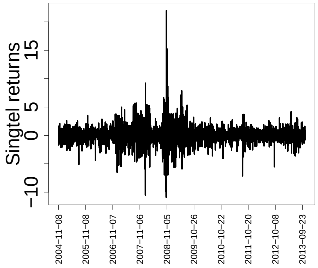
  where so far, we assumed $\varepsilon_t \sim N(0, 1)$.
- We can estimate model parameters through MLE and forecast $\sigma_{t|t-1}$ and construct prediction intervals for r_t.
- We also found that GARCH(p, q) processes can be written as ARMA$(\max(p, q), p)$ models for r_t^2, e.g. for GARCH(1,1):

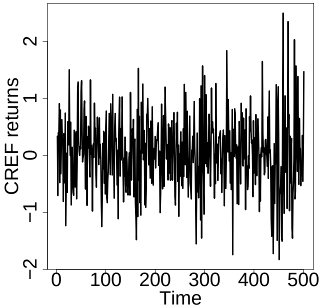
$$r_t^2 = \omega + (\beta_1 + \alpha_1)r^2_{t-1} + \eta_t - \beta_1\eta_{t-1},$$

  which is helpful for deriving some properties, e.g. the unconditional variance of r_t.

Next: let's analyze some real data!

- Steps:
  - Model identification:
    when to use a GARCH(p, q) model? How to choose p and q?
  - Model fitting and diagnostics checking
  - Forecasting
- Data: CREF, Singtel

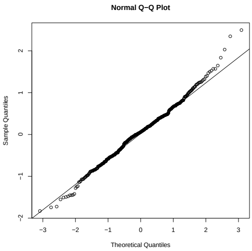


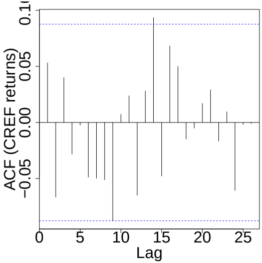


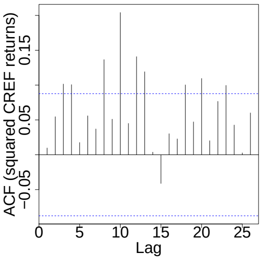
Model identification I

- When to consider using a GARCH(p, q) model?
- Example: for CREF returns, exploratory plots suggest
  - No autocorrelation in returns,
  - but autocorrelation in squared or absolute returns, combined with normality issues (both characteristics of GARCH processes).

  which indicates that a GARCH model may be appropriate.

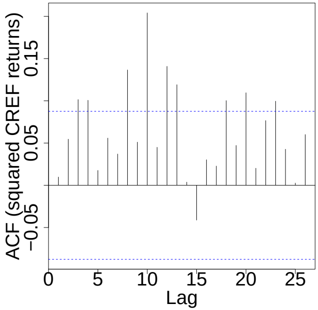


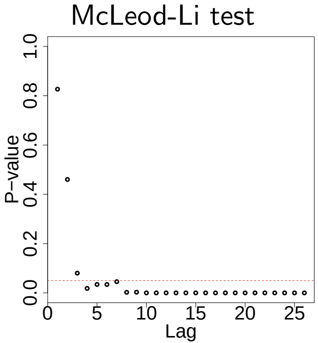




Investigating CREF return data: McLeod.Li test

- We can formally test whether there is 'conditional heteroscedasticity' (whether volatility changes with time) with the McLeod-Li test, which is the Box-Ljung test applied to squared returns.
  - H_0: no autocorrelation in squared residuals.
  - If we reject H_0, GARCH models can be considered.
- Conclusion for CREF data: H_0 is rejected (for higher lags) thus suggesting that there is conditional heteroscedasticity.





Model identification II: How to choose p and q?

- Remember that a GARCH(p, q) model corresponds to an ARMA($\max(p, q), p$) model for the squared returns:

$$\begin{aligned} r_t^2 &= \omega + (\beta_1 + \alpha_1) r_{t-1}^2 \\ &\quad + \ldots + (\beta_{max(p,q)} + \alpha_{max(p,q)}) r_{t-max(p,q)}^2 \\ &\quad + \eta_t - \beta_1 \eta_{t-1} - \ldots - \beta_p \eta_{t-p}, \end{aligned}$$

  where $\beta_k = 0$ for $k > p$ and $\alpha_k = 0$ for $k > q$.
- Book suggests examining ACF, PACF and EACF of squared or absolute residuals. However, in practice, this does not seem to work very well.
- GARCH(1,1) is often the default choice. If model diagnostics are not okay or if one of the coefficients is not significant, other models can be explored and compared with an information criterion.

Estimation

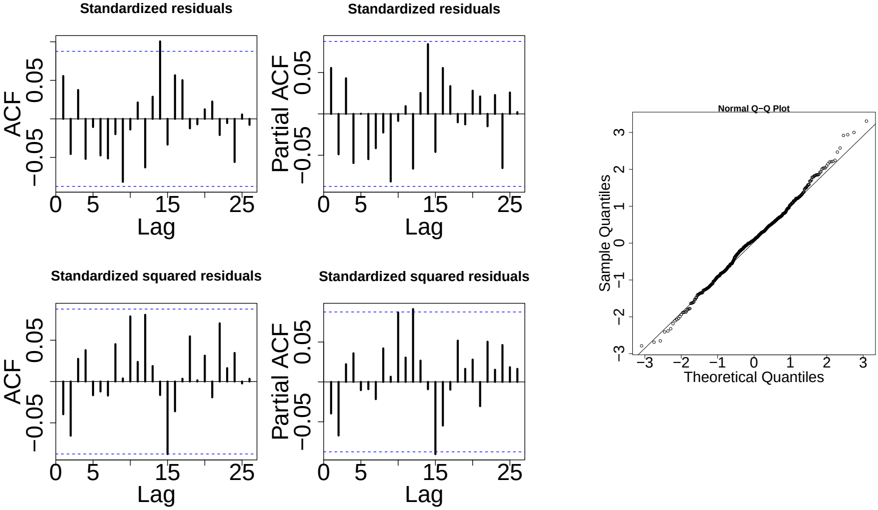
- As discussed, maximum likelihood estimation use “FitGarch”.
- Results for CREF:

```
s1 <- garchFit(formula = ~ garch(1, 1),
data = r.cref,  include.mean = FALSE)
summary(s1)
Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
omega    0.01644     0.01025    1.603   0.1088
alpha1   0.04423     0.01791    2.470   0.0135 *
beta1    0.91666     0.03514   26.089   <2e-16 ***
```

Model diagnostics

- Standardized residuals: $\hat{\varepsilon}_t = r_t/\hat{\sigma}_{t|t-1}$.
- Use plots and tests to verify
  - normality for standardized residuals,
  - lack of autocorrelation for squared standardized residuals.
- Book discussed tests in more detail (outside material), we use the standard output from the garchFit function.

Model diagnostics for CREF data: plots

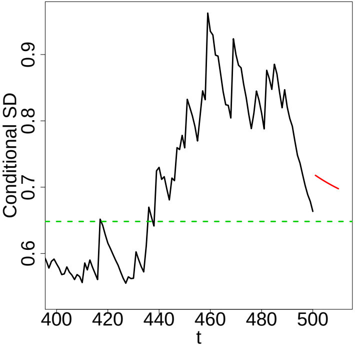


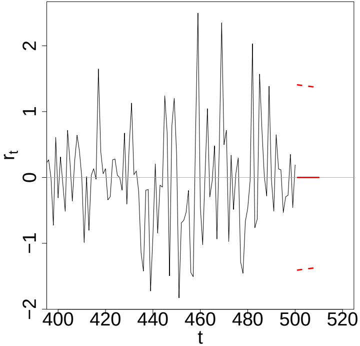
Model diagnostics for CREF data: tests

- See output fore various tests below:
  - First two tests (JB and SW) refer to testing normality for standardized residuals (H_0: normality),
  - LB tests refer to autocorrelation for standardized residuals R and squared standardized residuals R^2 (H_0: lack of autocorrelation).

```
> summary(s1)
...
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  1.110003  0.5740715
 Shapiro-Wilk Test  R    W      0.9964298 0.3298903
 Ljung-Box Test     R    Q(10)  10.9318   0.3628558
 Ljung-Box Test     R    Q(15)  19.42822  0.1949806
 Ljung-Box Test     R    Q(20)  22.57924  0.3099236
 Ljung-Box Test     R^2  Q(10)  8.65088   0.5655254
 Ljung-Box Test     R^2  Q(15)  16.60095  0.3432739
 Ljung-Box Test     R^2  Q(20)  19.34701  0.4993838
```

CREF volatility estimates and forecasts

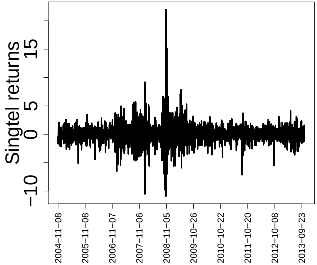


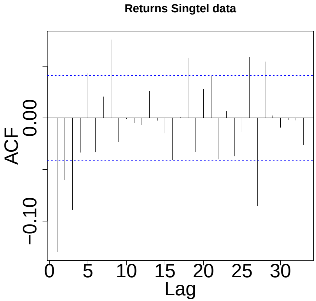


Summary

- Identification with plots and tests:
  Consider a GARCH model for non-autocorrelated time series
  - with autocorrelation in squared or absolute values,
  - where normality does not hold true.
- Fitting (MLE), diagnostics (tests and plots for the standardized residuals), forecasting (using standard formula):
  results obtained conveniently through built-in functions.
- What about Singtel data?

Singtel data: exploration





- Extend modeling approach discussed so far:
  - How to combine ARMA models and GARCH to model conditional mean and variance simultaneously?
  - How to overcome issues with non-normality of GARCH standardized residuals (innovations)?

Combined ARMA and GARCH modeling

- If the time series is autocorrelated, the GARCH model needs to be extended.
- The ARMA+GARCH model for a time series Y_t is given by

$$\left.\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \cdots + \phi_u Y_{t-u} \theta_0 + e_t + \theta_1 e_{t-1} + \cdots + \theta_v e_{t-v} \\ e_t &= \sigma_{t|t-1} \varepsilon_t \\ \sigma^2_{t|t-1} &= \omega + \alpha_1 e^2_{t-1} + \cdots + \alpha_q e^2_{t-q} + \beta_1 \sigma^2_{t-1|t-2} + \cdots + \beta_p \sigma^2_{t-p|t-p-1} \end{aligned}\right\}$$

  which is an ARMA process for Y_t, with (uncorrelated but dependent) white noise e_t modeled with a GARCH process.
- Modeling approach:
  - Ignore volatility in e_t to identify and fit an ARMA model using the standard approach.
  - Examine the residuals $\hat{e}_t$ of the ARMA model to identify whether a GARCH model would be appropriate (check autocorrelation and normality).
  - Fit a combined ARMA-GARCH model and check model diagnostics.

Singtel data Step 1: fit ARMA model

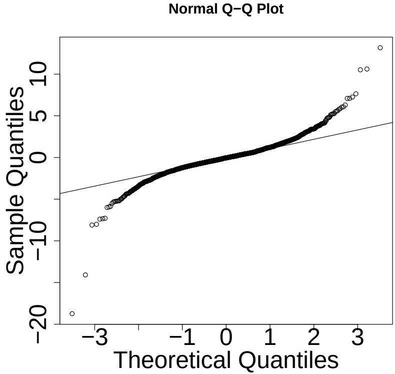
- MA(3) model for the returns

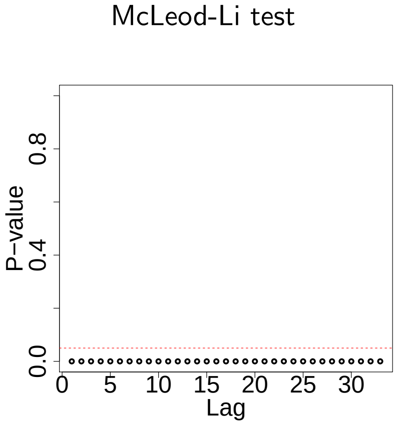
```
> auto.arima(returns,  ic = "bic",
 approximation = FALSE, stepwise = FALSE)


ARIMA(0,0,3) with zero mean

Coefficients:
          ma1      ma2      ma3
      -0.1571  -0.0696  -0.0950
s.e.   0.0207   0.0217   0.0214
```

Singtel diagnostics for residuals of MA(3) model





Fit a combined ARMA-GARCH model

- Fitting:

```
m2 <- garchFit(~arma(0,3)+garch(1,1),
  data = returns, include.mean = FALSE)
summary(m2)
Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
ma1     -0.07035     0.02268   -3.102  0.00192 **
ma2     -0.03799     0.02336   -1.626  0.10389
ma3     -0.07608     0.02322   -3.277  0.00105 **
omega    0.06143     0.01274    4.820 1.44e-06 ***
alpha1   0.10372     0.01391    7.457 8.86e-14 ***
beta1    0.87019     0.01614   53.926  < 2e-16 ***
```

- However... diagnostics are not ok...

Diagnostics for the ARMA-GARCH model

- Test results suggest that there are issues with normality:

```
> summary(m2)
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  672.5695  0
 Shapiro-Wilk Test  R    W      0.9802307 0
 Ljung-Box Test     R    Q(10)  7.802521  0.6481194
 Ljung-Box Test     R    Q(15)  10.10239  0.8132503
 Ljung-Box Test     R    Q(20)  12.9996   0.8774014
 Ljung-Box Test     R^2  Q(10)  3.914661  0.9511146
 Ljung-Box Test     R^2  Q(15)  5.191405  0.990378
 Ljung-Box Test     R^2  Q(20)  8.604858  0.9870485
```

- We can explore whether alternative model specifications are more appropriate (e.g. for the ARMA part, or the GARCH part) but it's not unlikely that innovations are not normally distributed.

Normality issues: a brief introduction

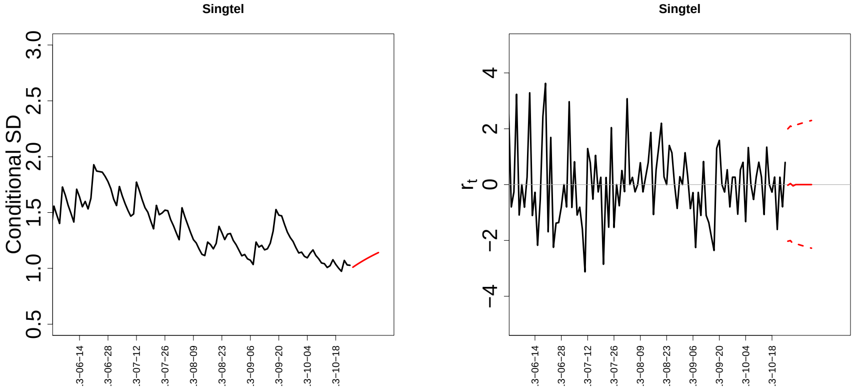
- Non-normality of innovations is not uncommon in financial time series.
- Several options:
  - Find a more appropriate distribution for the innovations ε_t and fit the model based on that distribution (outside class material).
  - Use quasi-maximum likelihood estimators (QMLE) to take into account that the distribution of the ε_t's is not normal.
- QMLE (details outside class material, just to introduce the main ideas):
  - Maximize the likelihood function based on assuming normality.
  - Estimators turn out to be approximately normal, centered at the truth (under mild conditions).
  - Adjust the SEs of the coefficients/forecast to account for non-normality.

QMLE in R for Singtel data

```
m3 <- garchFit(~arma(0,3)+garch(1,1),
data = returns, cond.dist = "QMLE", include.mean = FALSE)
...
Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
ma1     -0.07035     0.02383   -2.953  0.00315 **
ma2     -0.03799     0.02386   -1.592  0.11131
ma3     -0.07608     0.02481   -3.067  0.00216 **
omega    0.06143     0.02154    2.852  0.00434 **
alpha1   0.10372     0.02065    5.022 5.11e-07 ***
beta1    0.87019     0.02489   34.955  < 2e-16 ***
```

- Point estimates are the same but SEs have changed as compared to m2.
- No worries about results for normality tests; we still reject normality but now account for it in the SEs of the coefficients and in the forecasts.

CREF volatility estimates and forecasts



Summary

- The class of GARCH(p, q) models is used for estimating and forecasting volatility, which refers to the conditional variance or standard deviation $SD(r_t | r_{t-1}, r_{t-2}, \ldots)$ for some time series r_t (e.g. returns).
- We can consider using a GARCH model for non-autocorrelated time series
  - with autocorrelation in squared or absolute values,
  - where normality does not hold true.
- We discussed how to fit (MLE), diagnose (tests and plots for the standardized residuals) and forecast volatility and returns
  - for the easiest situation where the returns are not autocorrelated and a normal distribution seems appropriate for the innovations.
  - for more complicated situations, where returns are autocorrelated (such that an ARMA+GARCH model has to be used) and innovations may not be normal (QMLE estimation).