Hand-in 1

Michael Ståhle

February 2016

Exercise 2

a)

To investigate if there is any conditional heteroscedasticity in the data we plot the log returns against time.

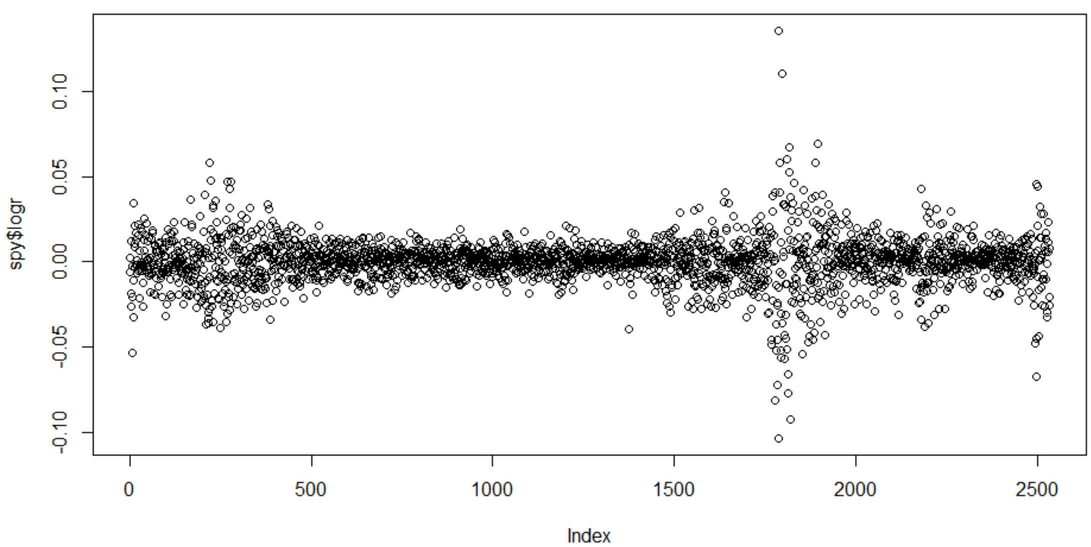

Figure 1: Plotted log returns against time.

Figure 1 clearly shows some variance clustering which is a sign of conditional heteroscedasticity.

Define r_t as the log return at time t, then serial correlation for lag l is defined as:

$$\rho_l = \frac{Cov(r_t, r_{t-l})}{Var(r_t)}$$

And the sample serial correlation for lag l is defined as:

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{T}(r_t - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=l}^{T}(r_t - \bar{r})^2}$$

Where T is the total number of time steps. Assuming r_t is a iid sequence satisfying $E(r_t^2) < \infty$, then $\hat{\rho}_l$ is asymptotically normal with mean zero and variance $1/T$ for any given l.

To jointly test if there is serial correlation present, a Ljung-Box test is preformed. We have the following hypothesis:

$$H_0 : \rho_1 = ... = \rho_m = 0$$

$$H_a : \rho_i \neq 0 \quad for\ any\ i$$

The test statistic is:

$$Q(m) = T(T+2)\sum_{l=1}^{m}\frac{\hat{\rho_l}^2}{T-l}$$

where $Q(m)$ follows a chi-squared distribution with m degrees of freedom. H_0 is rejected if $Q(m) > \chi_\alpha$ for a given α. Simulation studies suggests that choosing m to $m = ln(T)$ provides a better power performance. In our data set we have $T = 2535$, so $m = ln(2535) \approx 8$. Then the test statistic equals:

$$Q(m) = 39.5$$

which is highly significant. H_0 is rejected in favour for H_a, so serial correlation is present for some or all of the 8 first lags.

b)

First we look at the return series and the squared return series ACF and PACF where PACF is the partial auto correlation function.

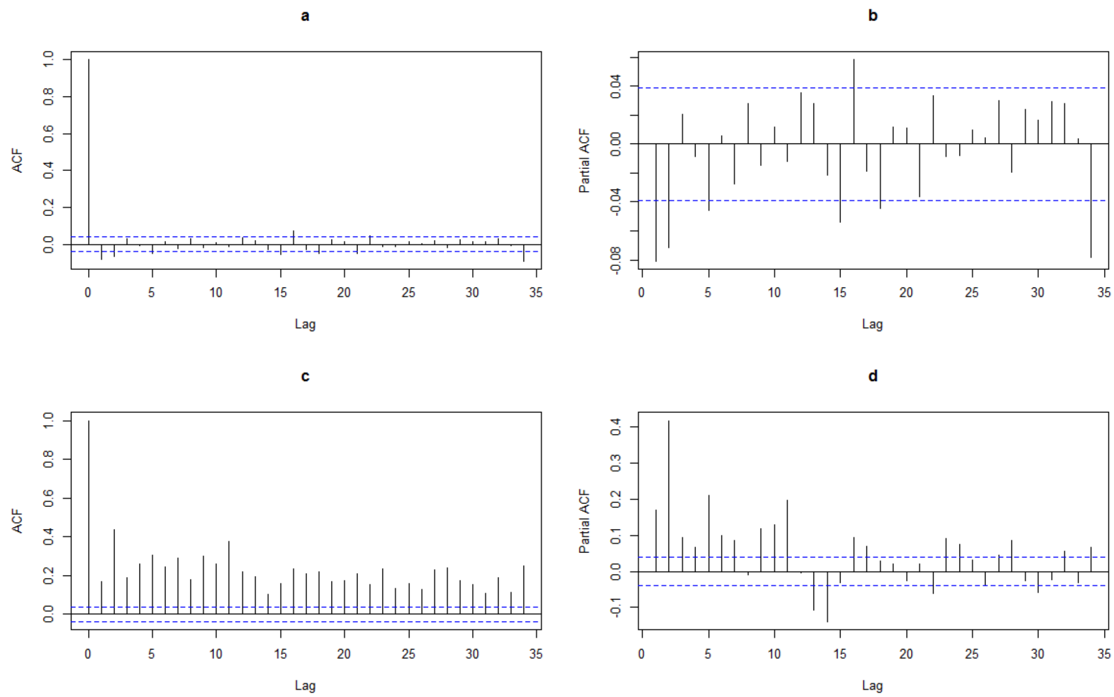

Figure 2: a and b shows the ACF and PACF for the log returns, c and d shows the ACF and PACF for the squared log returns.

According to figure 2, the log returns have most significant serial correlation for lag 1 and 2. The plots c and d shows a strong dependence between the squared log returns. After trying different combinations of ARMA models up to order 2, the best fit is an ARMA(2,0) model or equivalent an AR(2) model.

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t$$

The estimated coefficients for the AR(2) model is presented in figure 3.

```
Coefficient(s):
             Estimate  Std. Error  t value Pr(>|t|)
ar1        -8.645e-02   1.982e-02   -4.361 1.29e-05 ***
ar2        -7.158e-02   1.982e-02   -3.612 0.000304 ***
intercept   7.286e-05   2.354e-04    0.310 0.756927
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

sigma^2 estimated as: 0.000187
log likelihood:       7283.58
AIC Criterion:        -14559.16
```

Figure 3: Results of the estimated AR(2) model.

The AIC Criterion for the models with different orders where more or less the same but the residuals where less correlated in the AR(2) model and all coefficients where significant. The intercept is not significant different from zero which is to be expected looking at the log returns in figure 1. The estimated variance for a_t equals $\hat{\sigma}_a^2 = 0.000187$. The residual diagnostics is presented in figure 4.

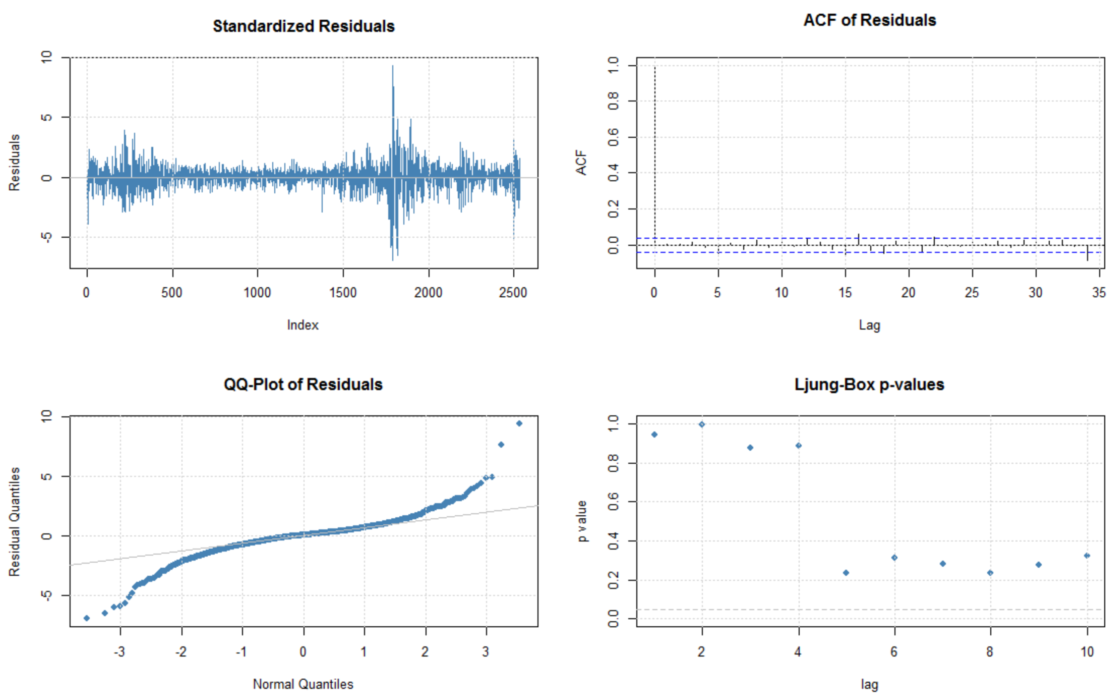

Figure 4: Residual diagnostics of the AR(2) model.

Figure 4 shows 4 different residual diagnostic plots. First we have the standardised residuals i.e. $\tilde{a}_t = \frac{a_t}{\sigma_a}$ plotted against time. It clearly shows that the residuals are heteroscedastic. Further the ACF and the plotted Ljung box p-values indicate no serial correlation in the residuals. Finally the QQ-plot shows that the residuals are not normal distributed. How ever, a PACF plot for the squared residuals shows a high dependence.

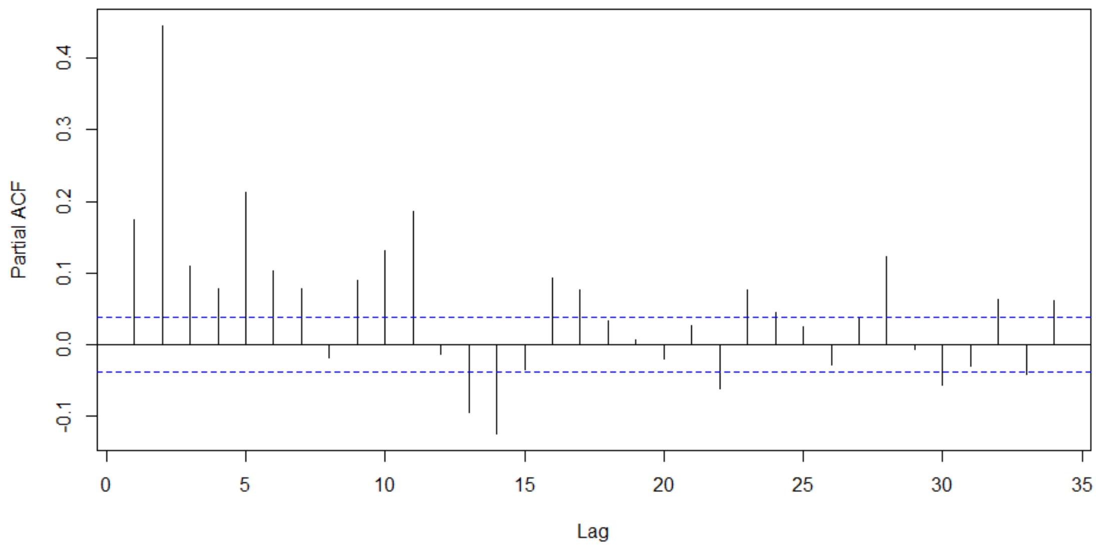

Figure 5: PACF of the squared residuals of the AR(2) model.

To model the dependence in a_t we want to fit a GARCH(m,s) model. After comparing models with different order, the model with the best fit is the GARCH(2,1) model.

$$a_t = \sigma_t \epsilon_t \quad \epsilon_i \sim iid \ and \ E[\epsilon_i] = 0$$

$$\sigma_t^2 = \omega_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \beta_1 \sigma_{t-1}^2$$

The result of the jointly estimated model i.e. AR(2)-GARCH(2,1) is presented in figure 6.

```
Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)    
mu     6.030e-04   1.704e-04    3.539 0.000401 ***
ar1   -6.643e-02   1.897e-02   -3.503 0.000460 ***
ar2   -3.233e-02   2.137e-02   -1.513 0.130321    
omega  2.337e-06   4.944e-07    4.727 2.28e-06 ***
alpha1 1.000e-08   1.201e-02    0.000 0.999999    
alpha2 1.136e-01   1.787e-02    6.358 2.04e-10 ***
beta1  8.701e-01   1.454e-02   59.851  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 6: Results of the estimated AR(2)-GARCH(2,1) model.

The coefficients for r_{t-2} (ϕ_2) and a_{t-1}^2 (α_1) are not significant, which suggests they should be dropped from the model. Figure 7 shows Ljung-Box tests for the residuals and the squared residuals.

```
Standardised Residuals Tests:
                                Statistic p-Value  
 Jarque-Bera Test   R    Chi^2  347.5857  0        
 Shapiro-Wilk Test  R    W      0.9835999 0        
 Ljung-Box Test     R    Q(10)  6.348361  0.7851957
 Ljung-Box Test     R    Q(15)  13.24688  0.5832357
 Ljung-Box Test     R    Q(20)  16.56222  0.6811903
 Ljung-Box Test     R^2  Q(10)  7.472945  0.6801678
 Ljung-Box Test     R^2  Q(15)  9.328941  0.8596997
 Ljung-Box Test     R^2  Q(20)  10.422    0.9598588
 LM Arch Test       R    TR^2   8.281886  0.7627322

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC 
-6.292995 -6.276875 -6.293010 -6.287147 
```

Figure 7: Tests for residuals and squared residuals for the AR(2)-GARCH(2,1) model

All of the Ljung-Box tests are insignificant. The AIC and BIC values are close to zero, indicating a god fit.

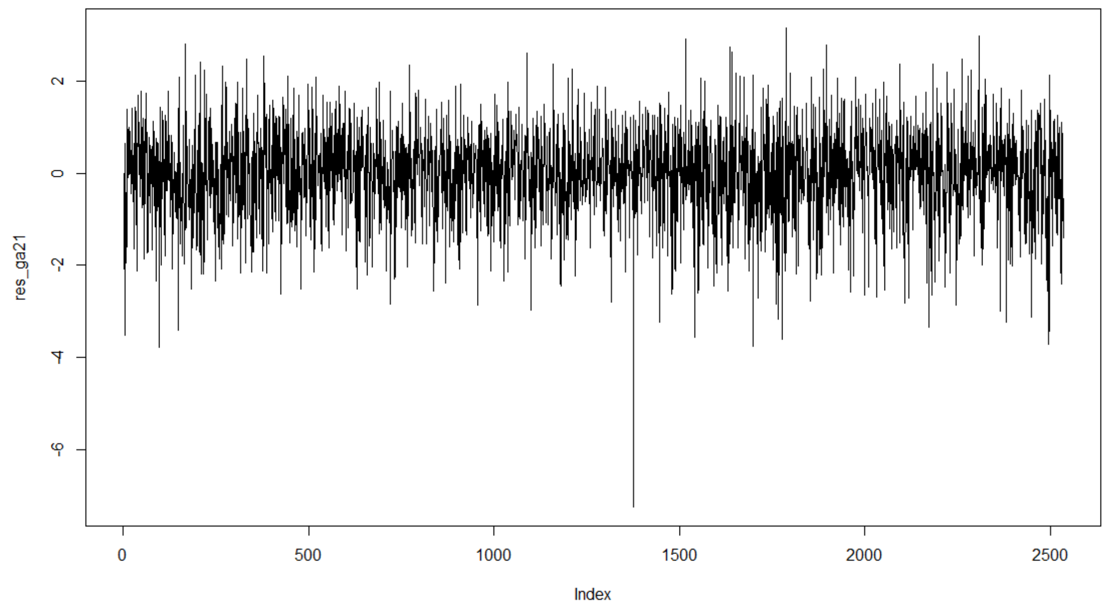

Figure 8: Standardized residuals plotted against time for the AR(2)-GARCH(2,1) model

Finally the standardized residuals are plotted in figure 8. They resemble a white noice serie. The proposed model AR(2)-GARCH(2,1) when dropping the insignificant coefficients is adequate.

c)

Previously we assumed that ϵ_t followed a standard normal. Now we assume that ϵ_t follows a student t distribution. The estimated coefficients are presented in figure 9.

```
Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)    
mu      7.245e-04   1.633e-04    4.437 9.14e-06 ***
ar1    -5.492e-02   1.847e-02   -2.973  0.00295 ** 
ar2    -3.874e-02   2.064e-02   -1.877  0.06046 .  
omega   1.584e-06   4.943e-07    3.204  0.00136 ** 
alpha1  1.698e-03   1.185e-02    0.143  0.88608    
alpha2  1.184e-01   2.004e-02    5.908 3.46e-09 ***
beta1   8.742e-01   1.609e-02   54.338  < 2e-16 ***
shape   7.643e+00   1.176e+00    6.497 8.22e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 9: Coefficients for the AR(2)-GARCH(2,1) model estimated with student t innovations.

The p-values for some coefficients becomes smaller and some becomes larger. There is no drastic change in the significance of the coefficients, although the coefficient for the ar2 effect is almost significant at the 5% level. Ljung-Box tests for the residuals and the standardized residual plot are almost identical to the ones in figure 7 and 8. The conclusion is that the estimation methods produce similar results.

d)

Now we want to fit an APARCH(s,m) model which is defined as follows:

$$r_t = \mu_t + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim D(0,1)$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^{s} \alpha_i (|a_{t-i}| - \gamma_i a_{t-i})^\delta + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^\delta$$

After testing combination of oreders up to (2,2), the AR(1)-APARCH(1,1) model has the most significant coefficients. The AIC values are marginally different between models with the different orders. The result of the AR(1)-APARCH(1,1) is presented bellow:

```
Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      4.761e-04  1.654e-04    2.878 0.004001 **
ar1    -4.801e-02  1.973e-02   -2.434 0.014927 *
omega   7.237e-05  1.902e-05    3.804 0.000142 ***
alpha1  6.381e-02  8.786e-03    7.263 3.80e-13 ***
gamma1  1.000e+00  1.692e-02   59.109  < 2e-16 ***
beta1   9.286e-01  8.088e-03  114.810  < 2e-16 ***
delta   1.164e+00  1.665e-01    6.993 2.69e-12 ***
shape   8.432e+00  1.393e+00    6.055 1.40e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 10: Coefficients for the AR(1)-APARCH(1,1) model estimated with student t innovations.

Figure 10 shows that all coefficients are significant at the 5% level. Residuals test are presented in figure 11.

```
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  710648.4  0
 Shapiro-Wilk Test  R    W      0.8533218 0
 Ljung-Box Test     R    Q(10)  10.70278  0.3811378
 Ljung-Box Test     R    Q(15)  14.45045  0.4916761
 Ljung-Box Test     R    Q(20)  16.81916  0.6646832
 Ljung-Box Test     R^2  Q(10)  13.91241  0.1770245
 Ljung-Box Test     R^2  Q(15)  13.93165  0.530719
 Ljung-Box Test     R^2  Q(20)  13.96304  0.8323652
 LM Arch Test       R    TR^2   15.82076  0.1995844

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-6.332069 -6.313646 -6.332089 -6.325385
```

Figure 11: Tests for residuals and squared residuals for the AR(1)-APARCH(1,1) model estimated with student t innovations.

All Ljung-Box tests are insignificant which means that the autocorrelation equals zero can't be rejected. The 5-step ahead forecast is presented in figure 12.

```
  meanForecast  meanError standardDeviation lowerInterval upperInterval
1 0.0016912044 0.02200316        0.02200316   -0.04223191    0.04561431
2 0.0003949390 0.02184914        0.02182358   -0.04322071    0.04401059
3 0.0004571777 0.02167199        0.02164658   -0.04280484    0.04371920
4 0.0004541894 0.02149732        0.02147212   -0.04245915    0.04336753
5 0.0004543328 0.02132516        0.02130016   -0.04211533    0.04302400
```

Figure 12: 5-step ahead forecast for the AR(1)-APARCH(1,1) model

e)

The qq-plots for the three models are presented in Figure 13-15. From the qq-plot in Figure 13 we observe that the left tail of the sample distribution is heavier than the left tail of the models theoretical distribution. When estimating the AR(2)-GARCH(2,1) model with student t innovation, we allow for heavier tails.

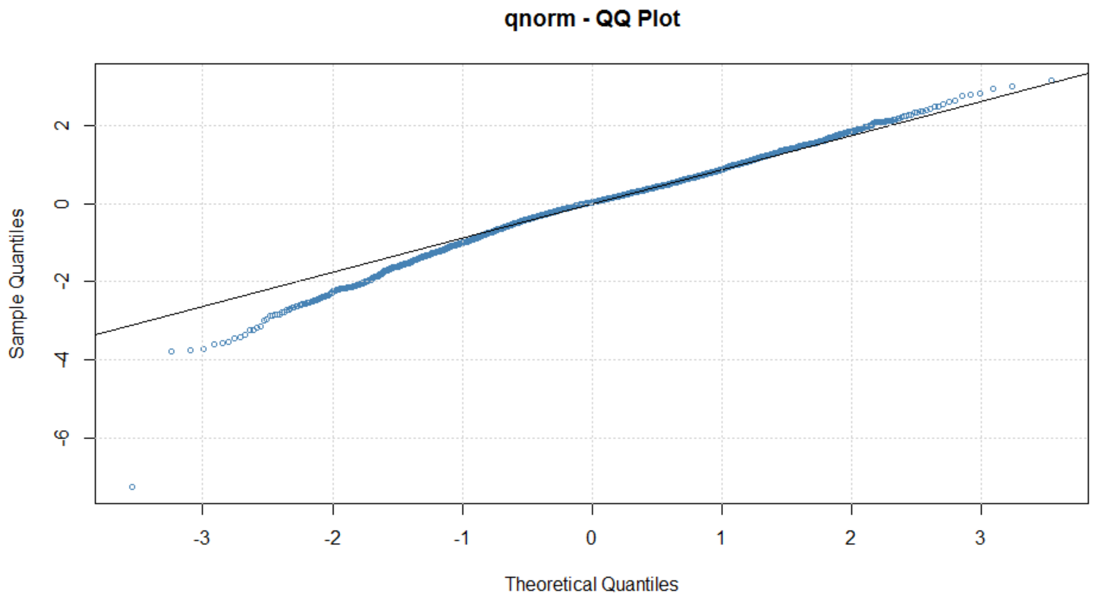

Figure 13: qq-plot for AR(2)-GARCH(2,1) with normal assumption

Figure 14 shows the qq-plot of the standardized residuals of the AR(2)-GARCH(2,1) model with student t innovations. We observe that the left tail has a better fit than in figure 13 but now the right tail is heavier for the theoretical model contra the sample. The difference implies that distribution for the standardized residuals are asymmetric.

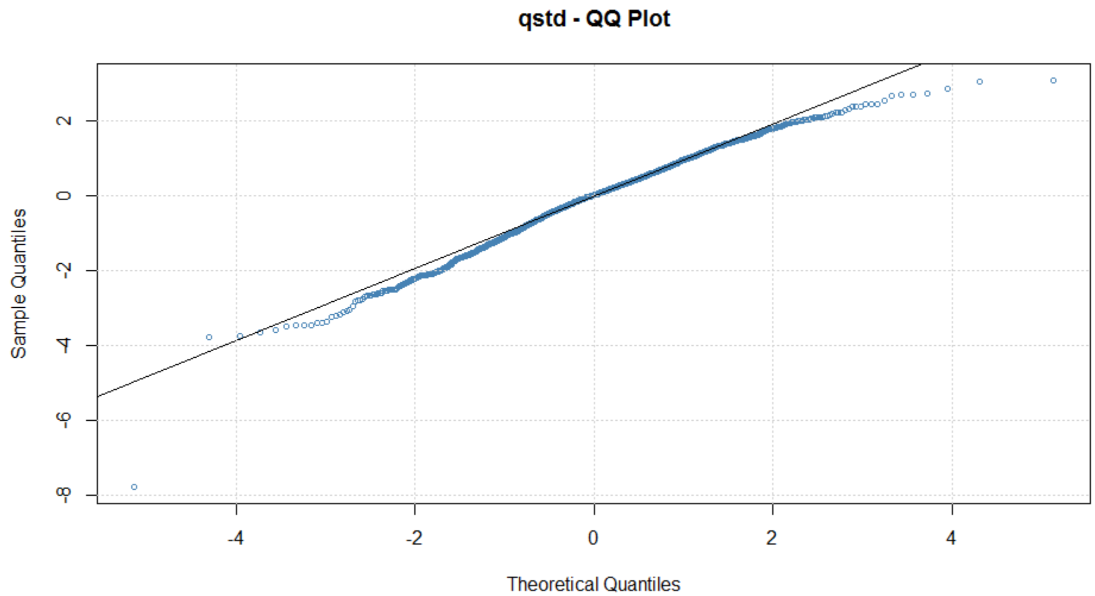

Figure 14: qq-plot for AR(2)-GARCH(2,1) with student t innovation

Finally, Figure 15 shows the qq-plot of the standardized residuals of the AR(1)-APARCH(m,s) model. The APARCH(m,s) model allows for asymmetry in return volatility. Except the three most extreme observations to the left, both tails of the sample fits the theoretical distribution well. As the other model diagnostics for the model were adequate, the AR(1)-APARCH(m,s) appears to be the best. Note that the most extreme residual deviation might be an outlier since it was present in all the qq-plots. It should might be removed from the data set.

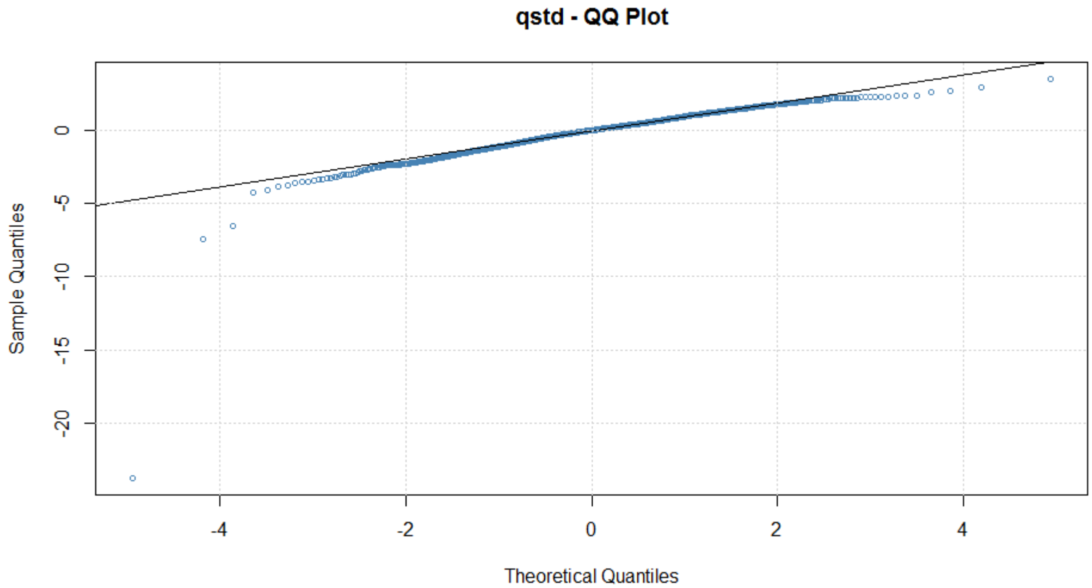

Figure 15: qq-plot for AR(1)-APARCH(1,1) with student t innovation