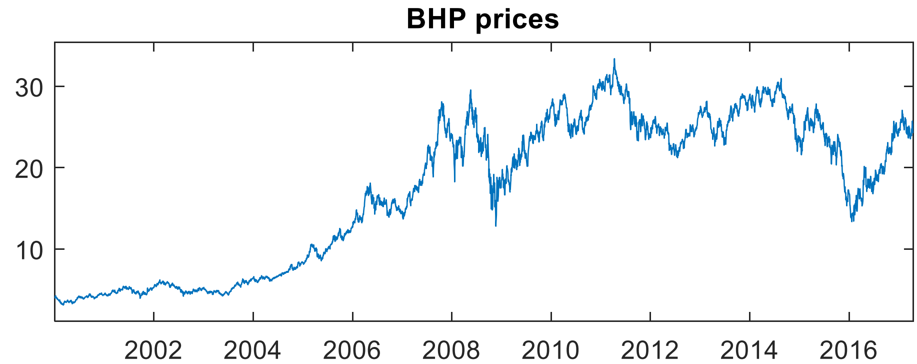
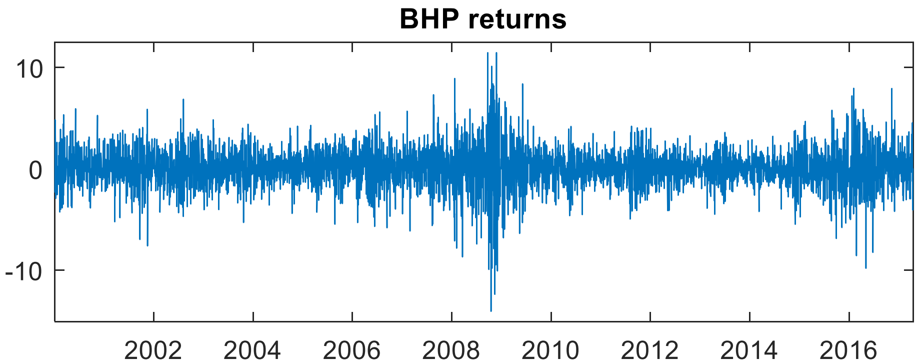
QBUS6830 Financial Time Series and Forecasting
S1, 2019

Solutions to Lab Sheet 8: ARCH Models

Q1 (ARCH model)
The dataset BHP00-17.csv contains daily prices on BHP stock from 04/01/2000 to 17/04/2017. Use Matlab and the Econometrics toolbox to answer these questions.

(a) Transform the prices to percentage log-returns. Plot the price and return series. Comment.





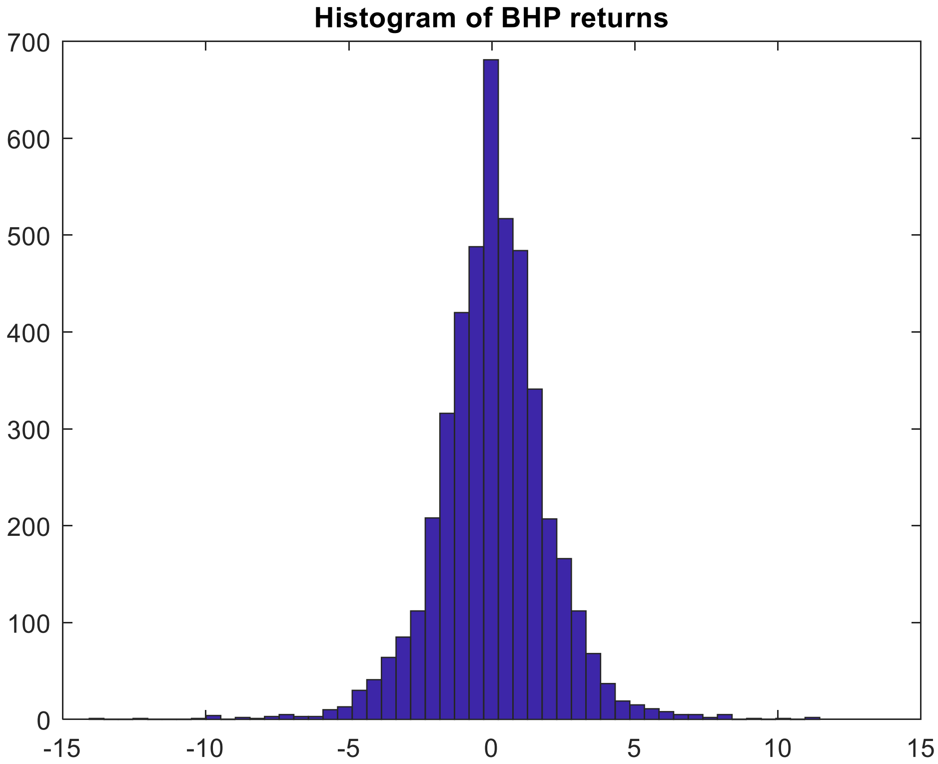
(b) Calculate some relevant summary statistics and then comment on the performance of BHP stock over the sample period. Perform the JB test for Gaussianity on the return series.

Mean	Median	Std dev	Skewness	Kurtosis	Min	Max
0.0397	0	1.9350	-0.1562	6.6719	-14.0772	11.4645

Percentiles

0.1	0.5	1	10	90	99	99.5	99.9
-6.1493	-4.9422	-2.1341	-1.0267	1.1166	2.2877	5.1615	6.2152

BHP stock has risen by $37.29, from around $4.00 in January, 2000 to $24.31 in April, 2017, which is an increase of 507% in value over nearly 16.5 years. Over that time, the biggest single one-day loss was 14.1%, in September, 2008, during the financial crisis. BHP did very well, with steady price growth, up until end of 2007 when it fell. It rose again at the start of 2008, peaked in May, 2008 then fell sharply and lost half of its value by the end of November, 2008. The stock then experienced solid gains for a couple of years until reaching a peak in April, 2011, followed by a price stagnation for a few years. In mid to late 2015, a sharp drop in prices was observed, followed by a recovery to near its peak level. Clearly, the Global Financial Crisis (GFC) period had higher variance, as the price dropped sharply by more than 50% in about 6 months. The other volatile period is associated with the ending of mining boom in 2015 which again saw the depreciation of 50% of its value.



There are several outlying single day returns over this period. Two negative ones stand out.

The kurtosis is 6.67 and skewness is -0.16. Clearly the tails appear fatter than a Gaussian, though the distribution appears symmetric: two large negative outliers have likely caused the negative skewness estimate. The JB test reveals that the p-value is < 0.001, thus the null hypothesis of Gaussian data (actually of skewness = 0 and kurtosis = 3) can be rejected.

(c) Fit an ARCH(1) model to the return data using ML. Report and interpret the parameter estimates and plot the estimated volatility series. Discuss the estimated series and its' properties.

Maximum likelihood output is

```
                              Standard          t     
 Parameter        Value          Error       Statistic
-----------    -----------   ------------   -----------
```

Constant	2.76564	0.0544427	50.7991
ARCH{1}	0.25839	0.0159109	16.2398
Offset	0.0508529	0.026067	1.95083

(Note that on assignments I expect MUCH better presentation than this table)

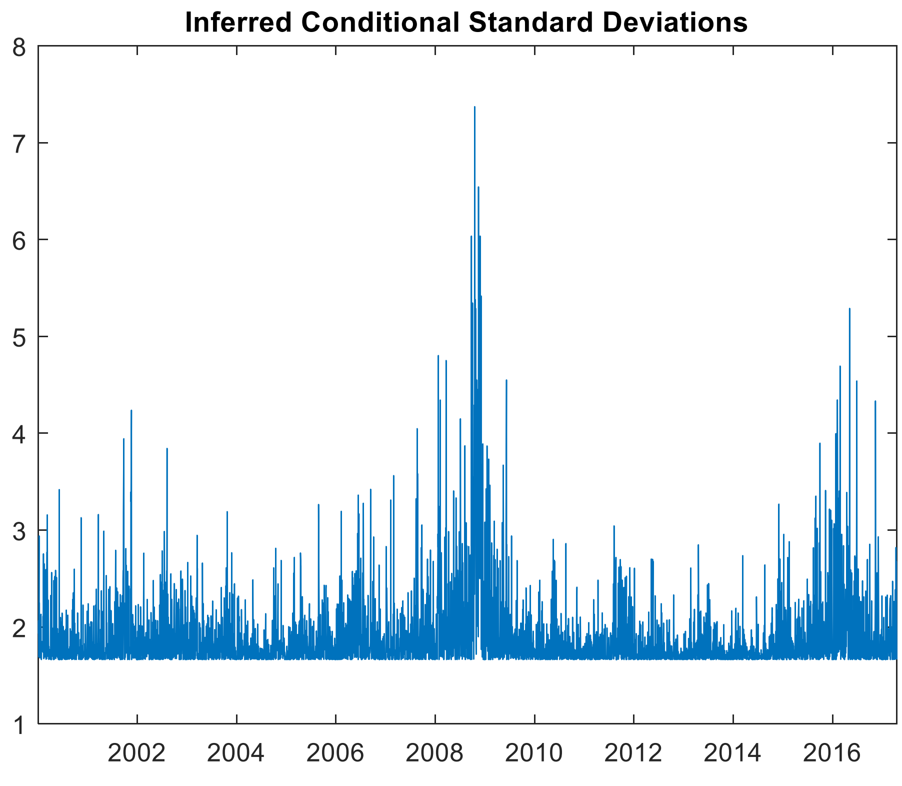
Thus the estimated model is:

$$r_t = 0.051 + a_t;\ a_t|I_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = 2.765 + 0.258 a_{t-1}^2$$

The estimated long-run (unconditional) mean daily return is 0.051% and is significantly different to 0 (t-stat 1.95). The estimated volatility following a return shock of $a_{t-1} = 0$ (i.e. return = mean return) is 2.765 (percent squared). The impact of previous shocks is that, if yesterdays squared shock had been higher by 1 (percent squared), today's volatility would be 0.258 (percent squared) higher, on average.

The volatility series is plotted below:



The series of conditional standard deviations (i.e. $\hat{\sigma}_t = \sqrt{Var\left(r_t \mid \Im_{t-1}\right)}$) seems very non-smooth and also to (strangely) have a lower bound that the standard deviation estimates cannot go below, but can hit. Why is this?

Clearly volatility increased for BHP in September 2007and continued over the GFC period, but had recovered to normal levels by mid-2009 or so. The end of the mining boom in mid to late 2015 saw another wave of high volatility which continued into 2016.

The volatility estimates seem to change starkly from day to day, sometimes by more than 4 or 5%! Many analysts don't believe this could really happen in true volatility (as often as it seems to in the estimated volatility above).

(d) Compare the unconditional variance estimated from the ARCH model with the sample variance? Are they close?

$$\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1} = 3.7292; \; s^2 = 3.7441$$

Yes, quite close.

(e) Compare the unconditional kurtosis estimated from the ARCH model with the sample kurtosis? Are these close?

$$\frac{3(1-\hat{\alpha}_1^2)}{1-3\hat{\alpha}_1^2} = 3.5009; \; \widehat{K} = 6.6719$$

No, these are not at all close to each other.

(f) Find the LS (using the squared errors) estimates for the parameters. Are these close to the ML estimates?

From the lecture notes:

For LS estimation, simply fit an OLS regression of the squared errors (a_t^2) on their lagged values (a_{t-1}^2). See matlab code for details. The OLS estimates are:

$$\hat{\alpha}_{0,LS} = 2.8819; \; \hat{\alpha}_{1,LS} = 0.2299$$

which are reasonably close to the ML estimates of

$$\hat{\alpha}_0 = 2.765; \; \hat{\alpha}_1 = 0.258$$

(g) Repeat parts (d) and (e) for the LS estimates.

$$\frac{\hat{\alpha}_{0,LS}}{1-\hat{\alpha}_{1,LS}} = 3.7424$$
$$\frac{3(1-\hat{\alpha}_{1,LS}^2)}{1-3\hat{\alpha}_{1,LS}^2} = 3.3770$$

The unconditional variance estimate is very similar to that for the ML estimates and the sample variance. The estimated kurtosis is smaller than that for the ML estimates and again nowhere near the sample kurtosis.

(h) Are the unconditional variance and kurtosis estimates from the LS and ML estimates close to each other? Are they close to the sample variance and kurtosis? If they are not, why do you think this might be?

The LS and ML estimates give variance and kurtosis estimates that are reasonably close to each other. The estimates of variance by both methods are close to the sample variance of BHP returns.

However, both ML and LS estimation methods give estimates of kurtosis that are well below the sample kurtosis estimate of 6.67. This is likely because the sample kurtosis is highly influenced by, and sensitive to, outliers in the data. Recall that the sample kurtosis is the expected value of the data points minus their mean all put the power of 4. Outliers are going to dominate this sample estimate.

The ARCH model has a way of accounting for outliers, by assigning them an increased variance, σ_t^2. The sample kurtosis cannot do that, it treats (i.e. weights) all observations equally, in their influence on sample kurtosis (and variance and mean too).

As such, I trust the model estimates more than the sample kurtosis.