Trevor Mitchell

Times Series Forecasting – Stat 2302

Project 2

The data I have chosen to build an ARIMA-ARCH model for is the effective federal funds rate from 1954-07-07 to 2019-12-11. Here is the time series plot of the data:

A screenshot of a cell phone

Description automatically generated

**Observations and Model Building**

In order to check for any level-dependent volatility in the dataset I decided to take logs. After taking logs I was able to remove a great deal of level-dependent volatility.

A screenshot of a social media post

Description automatically generated

After taking logs, however, there did appear to still be a trend component in the data. The ACF of the original time series dies down very slowly, so it seemed very appropriate to difference the data.

**A screenshot of a social media post

Description automatically generated A screenshot of a social media post

Description automatically generated**

After differencing the data, the first lag of the ACF of the differenced data became negative. However, it was only -.201239 as opposed to the original .995763 of the log federal funds rate dataset. I also checked the 1st lag of the second difference which was -.533977, so I reasoned that differencing twice was too much. In addition, the first difference’s ACF plot goes to zero much quicker than the dataset before differencing and yields the smallest standard deviation with value of 0.11897 compared to values of 1.3124 (Original Series) and 0.18440 (2nd differenced series).

A screenshot of a cell phone

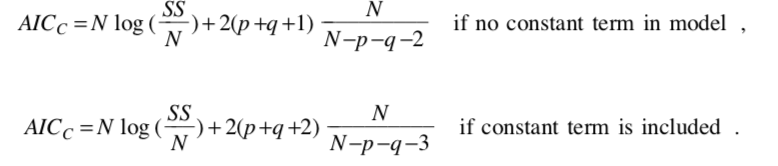
Description automatically generated

The first difference also yields a time series that fluctuates around a well-defined mean. Differencing the series once appears to make it stationary.

A screenshot of a cell phone

Description automatically generated

After figuring out how much differencing my data needed, I used an AIC\_C grid search to find the optimal parameters for my ARIMA model. I allowed AIC\_C to determine whether or not a constant was needed in my model. As a result of my AIC\_C grid search, I found ARIMA\_414c (with constant) to be my optimal model. I used the respective AIC\_C formula for models that do and do not include a constant term.



A close up of text on a white background

Description automatically generated

A close up of text on a white background

Description automatically generated

**A screenshot of a cell phone

Description automatically generated**

**Complete form of Fitted Model (ARMA)**

A close up of a logo

Description automatically generated

**Ljung-Box Statistics**

The Ljung Box statistics seemed to indicate that my ARIMA model is not adequate, as my p-values appear to be zero which is much less .05. However, the ACF and PACF plot of the residuals seem to indicate that my model chosen is okay. I have autocorrelations and partial autocorrelations near zero for most lags. Most of the lags that are a bit farther from zero are not statistically significant and those that are significant are barely so. The plot itself also appears to have zero mean.

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated

**Forecasts of ARIMA\_012 at Lead Times 1-50**

The forecast interval appears to be a bit too narrow. Much of the historical datapoints fall outside of the forecasts. However, recent data points do appear to be within the forecast interval.

A screenshot of a social media post

Description automatically generated

ARCH Component

Plots of Residuals Squared Time Series, ACF, & PACF

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated A screenshot of a social media post

Description automatically generated

While the residuals are approximately uncorrelated, the ACF and PACF of the squared residuals have significant lags which is reasonable evidence that the residuals are not independent and instead suggest evidence of conditional heteroscedasticity.

According to AIC\_C the optimal ARCH(q) model is the ARCH(10). The parameter values for a0, a1, a2, a4, a5, a6,a7, and a10 are quite statistically significant. While a9 is significant at .01 level of significance, a8 is significant at .05, and a3 is not statistically significant. The complete form of the ARCH(10) model is as follows below:

A close up of a logo

Description automatically generated

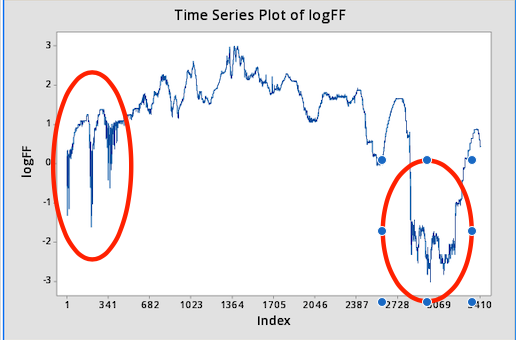
A screenshot of a cell phone

Description automatically generated

The unconditional variance of the shocks was computed to be 0.014975856230972646 by taking the expectation.

The 95% one-step ahead forecast interval for the ARIMA\_ARCH model is (0.281189, 0.608628). The 5th percentile of the conditional distribution is 0.30708375783198616 for the next periods log federal funds rate.

As one can see from the plots below, the highly volatile periods for the log federal funds rate appear to coincide with the highly volatile periods of the conditional variances plot

 A picture containing screenshot

Description automatically generated

The forecast intervals appear to be very useful and practical as the log federal funds rate appears to fall relatively tightly in between the upper and lower bounds of the ARIMA-ARCH intervals.

A screenshot of a cell phone

Description automatically generated

In accordance with the normality tests of my ARIMA-ARCH residuals, my model does not seem to follow a normal distribution as I have a p-value less than .005. The model does not seem to adequately describe “long-tailedness” in the data. The model appears to show increasing departure from the fitted line above and below for data points on the end.

A close up of a map

Description automatically generated

The ARIMA-ARCH intervals failed 169 times which accounted for 4.95021 % of the time.

The one-step ahead forecast for the ARIMA model and the ARIMA-ARCH model both contained the actual target value of 0.438255 (the final omitted data point) with ARIMA having an interval of (0.228941, 0.66088) and ARIMA-ARCH having an interval (0.281189, 0.608628). While both intervals worked, the ARIMA-ARCH model had a tighter upper and lower bound on the target. It would seem that the ARIMA-ARCH generated a superior prediction interval.

**Source of the Data Set**

<https://fred.stlouisfed.org/series/FF>