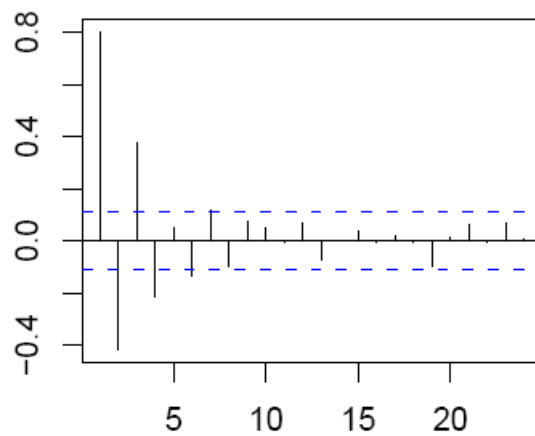
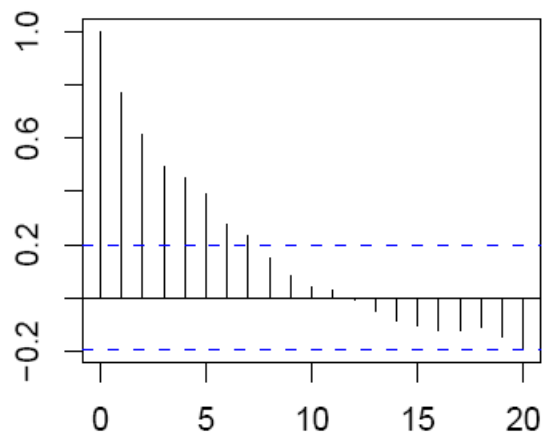


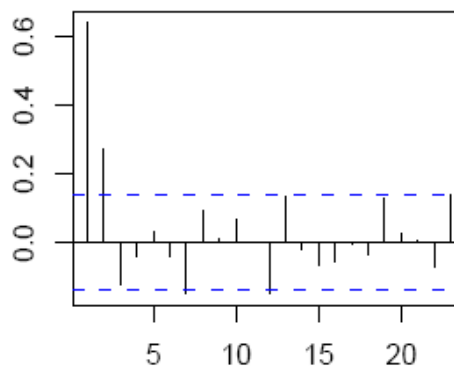
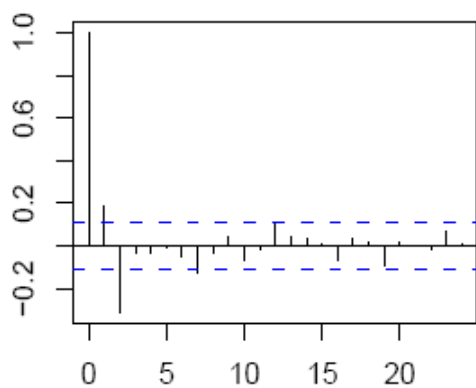
Time Series Graphs

Model	ACF	PACF
White Noise	All zeros	All zeros
AR(p)	Exponential Decay	P significant lags before dropping to zero
MA(q)	q significant lags before dropping to zero	Exponential Decay
ARMA(p,q)	Decay after q^{th} lag	Decay after p^{th} lag

Examples of Exponential Decay:



Examples of Significant Lags:



Time Series Analysis Procedure

1. Stationarity

The Box-Jenkins ARIMA modelling that we use can only be applied to time series that are **stationary**. The first step in the analysis is therefore to make our series stationary.

- i. Time Series Plot – constant mean and variance implies stationarity
- ii. Trend Stationary – remove a trend/slope and the series that is left is stationary
- iii. Difference Stationary – difference the series and it becomes stationary

2. ACF & PACF

Once you have found the stationary form of the series you look at the ACF and PACF plots to determine the form of the ARIMA(p,d,q) model for the stationary series.

3. Fit the model

The model you decide on is what we will call the “**Original Model**”

Fit this model to the stationary series.

The output in R may be used to indicate whether or not specific terms are needed in the model.

The null hypothesis is that the coefficient of the term is equal to zero. Therefore, a low p-value implies that the coefficient is significantly different from zero and the term is needed in the model.

4. Residual Analysis

There are three sections in residual analysis:

- i. Residual Plots
- ii. Overfitting
- iii. Ljung-Box Pierce Statistic

i. Residual Plots

Each of the five plots in your residual analysis gives you a different piece of information about the residuals of your model.

Remember for your model to be a good fit you want your residuals to be **white noise** terms (i.e. $\epsilon_t \sim \text{iid } N(0,1)$).

- a. ACF plot of residuals – this plot shows whether or not there is autocorrelation present in the residuals.
- b. Normal (or QQ) plot of residuals – this plot tests for Normality in the residuals.
- c. Histogram of residuals – this plot shows the distribution and spread of the residuals.
- d. Residuals vs. the fitted values – shows trends and relationships within the residuals.
- e. Residuals vs. the Order – shows mean and variance of residuals.

ii. Overfitting

Overfitting is used to make sure that you have not left any terms out of the model. You start with your **Original Model** and you add terms to it.

Step 1:

Write down the form of your **Original Model**

Step 2:

Fit an extra **AR term** to the **Original Model**

Look at the p-value to figure out if this term is needed or not

If not, move on to **Step 3**

If the term is needed, you now have a new form for your Original Model – go back to **Step 1**

Step 3:

Fit an extra **MA term** to the **Original Model**

Look at the p-value to figure out if this term is needed or not

If not, move on to **Step 4**

If the term is needed, you now have a new form for your Original Model – go back to **Step 1**

Step 4:

Write down the new form of your model

This is your **Final Model**

There are a few other things to notice during Overfitting:

- Sometimes, when you overfit a term it will knock one of the original terms out of the model. When this happens, look at the p-values and fit the terms that are needed. You then compare the fit of this model to the fit of your original model by looking at the residual plots and at the estimate of the variance i.e. the MSE.

Decide on which model is the best and go back to Step 1 in Overfitting with this new Original Model.

- As a rule of thumb fit AR terms first.

One suggested reason for this is that the AR terms are simply lagged Y terms. They hold a lot of information about the series. The MA terms are calculated from the model and may hold less information ... think about it.

When the Overfitting gives you two models to choose from it is actually telling you that these two combinations of terms hold approximately the same information on the series.

In general, one AR term will often explain more than one MA term. By the law of Parsimony, we want to have the least number of terms in our model. Therefore, we look at the AR terms first.

iii. Ljung-Box Pierce Statistic

This is sometimes referred to as the 'Modified Box-Pierce (Ljung-Box) Chi-Square statistic'. This statistic is often quoted in research articles and so it is important to know ... but PLEASE NOTE that the R graph of Ljung-Box-Pierce p-values is not a useful graph in most cases!!

The null hypothesis for this test is that your model is a good fit.
Therefore, **you want to fail to reject** the null hypothesis.

To do this you need **large p-values** ($p > \alpha$: Fail to reject H_0). The test gives you Ljung-Box p-values at different lags –all of these p-values should be large.

Remember: There are theoretical issues with the Ljung-Box Pierce statistic. It should be the last thing you check. If all other residual analysis indicates that your model is a good fit, don't worry too much if the Ljung-Box p-values are too small.

5. Final Model

Once you have finished the residual analysis, write down the final ARIMA(p,d,q) form of the model.

Don't forget

It is the original series that you are writing the model for. If you had to adjust it to make it stationary you need to take that into account in the final form of the model.