SELECTING A TIME SERIES MODEL

We consider a systematic approach to analyzing time series data using the following example. Here is a basic summary of the data.

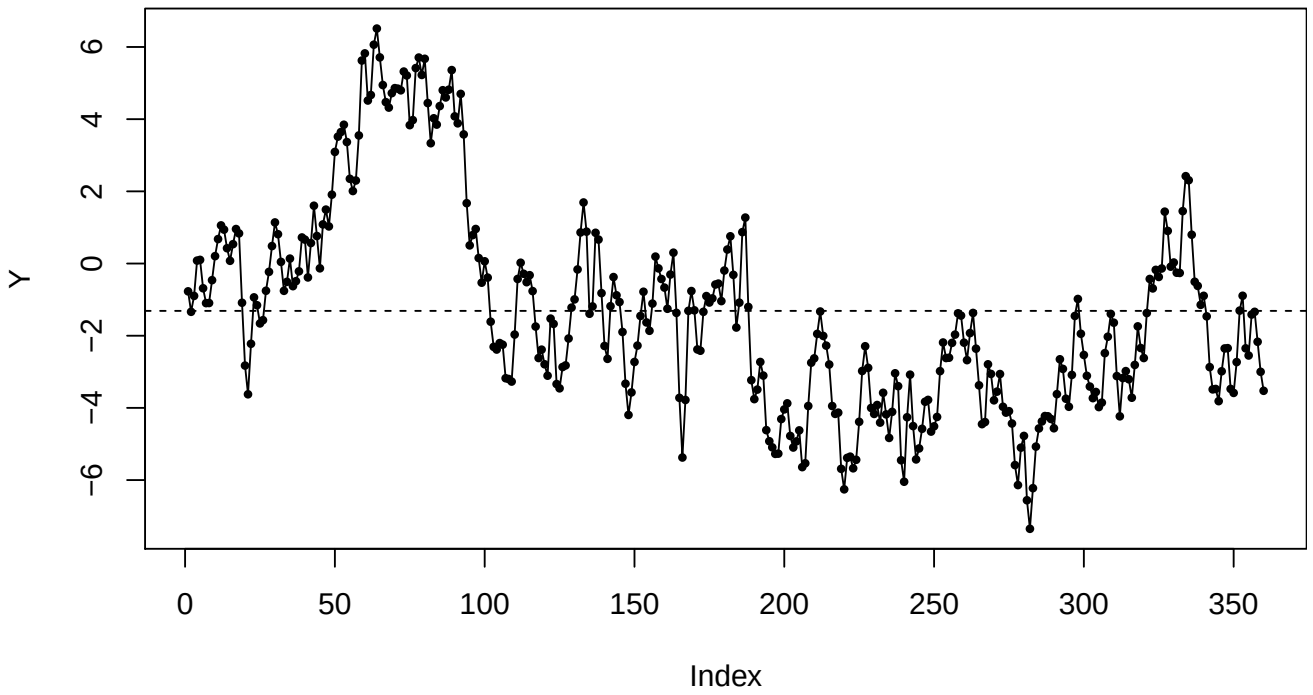
```
length(Y)

+ [1] 360

Y[1:20]

+  [1] -0.77065921 -1.33698499 -0.89716653  0.08133360  0.10167698
+  [6] -0.68518330 -1.09666052 -1.09380375 -0.46023937  0.20449318
+ [11]  0.67818563  1.05737545  0.93970306  0.42467640  0.07601754
+ [16]  0.53716565  0.95579078  0.83576087 -1.08645763 -2.82731431

par(mar=c(4,4,1,0))
plot(Y,pch=19,cex=0.5);lines(Y)
abline(h=mean(Y),lty=2)
```



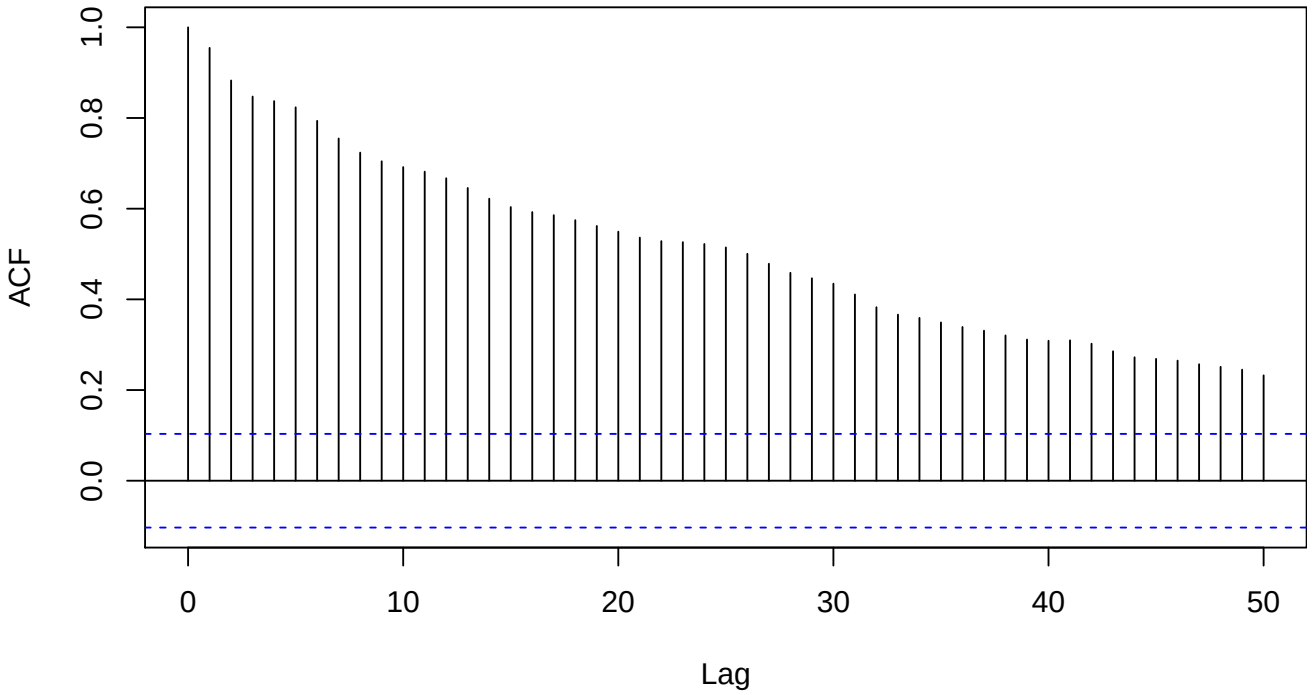
There is no obvious systematic trend or periodicity, as indicated by the deviations about the plotted mean.

```
mean(Y)

+ [1] -1.312014
```

However, there does seem to be strong positive dependence in the series. This can be assessed by inspecting the sample acf:

```
par(mar=c(4,4,1,0))
acf(Y,lag.max=50)
```

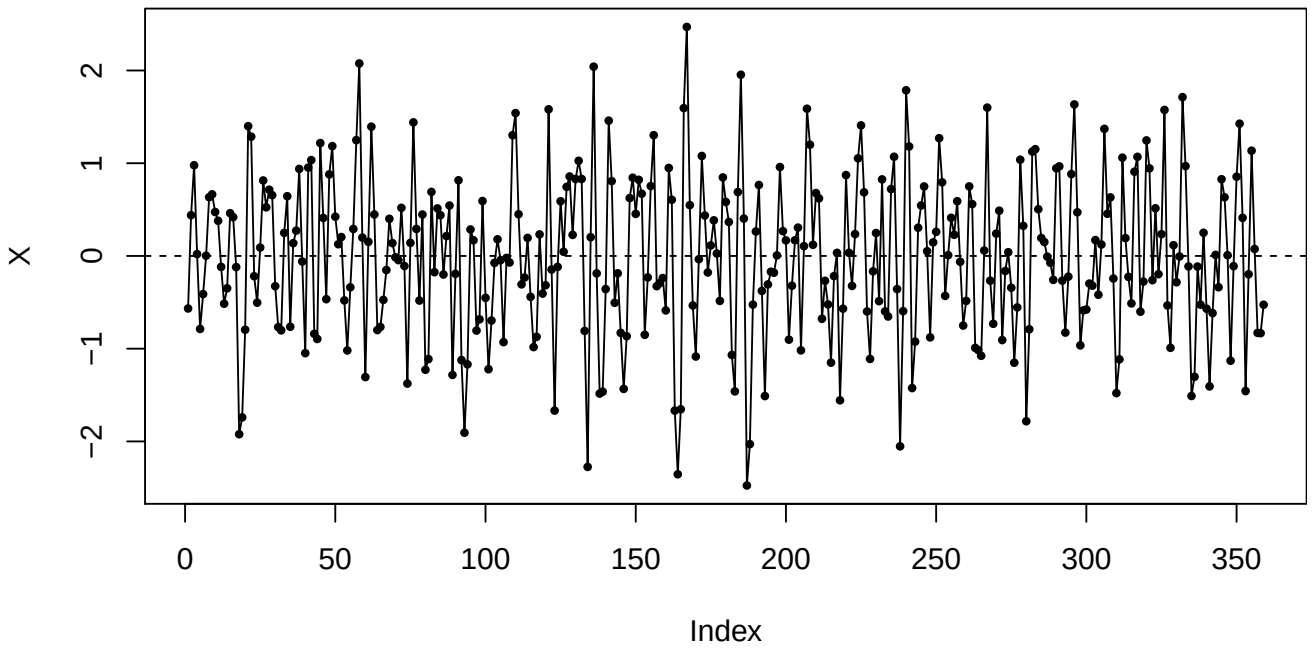


This plot (with a slow decay of the acf) indicates that a form of non-stationarity may be present. Recall that there are different ways that non-stationarity can be present:

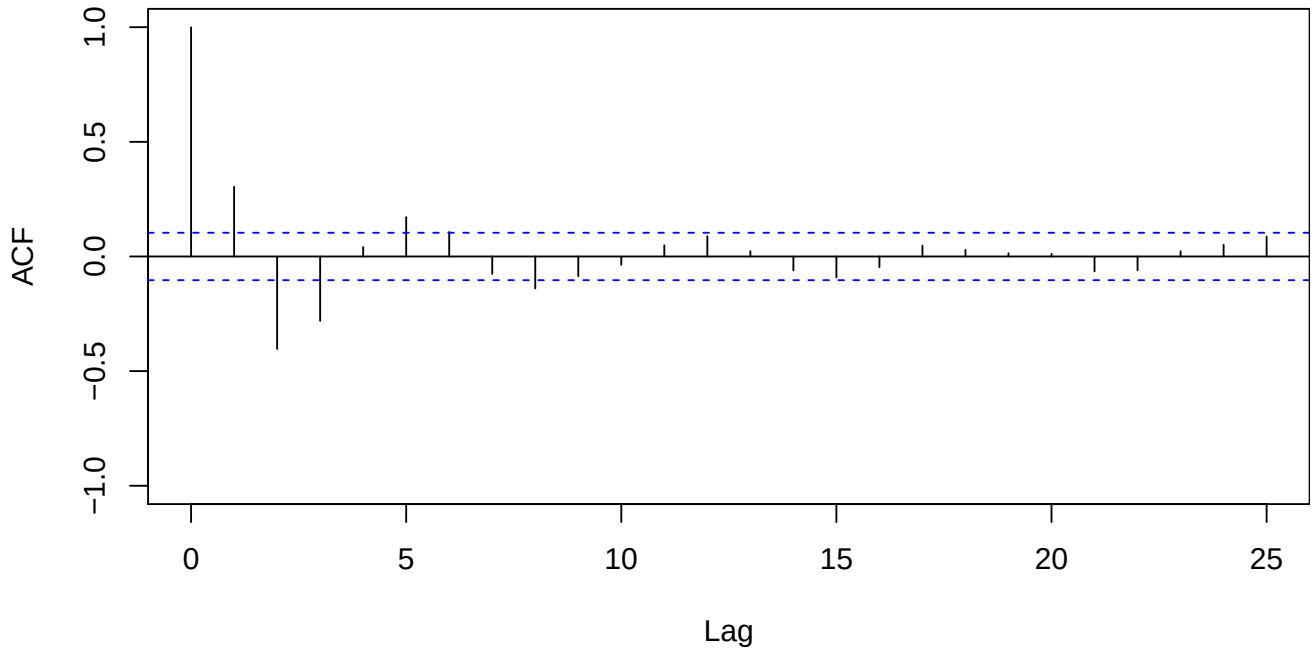
- deterministic variation in mean or covariance;
- deterministic periodic or seasonal component;
- stochastic non-stationarity due to a *unit root* in an AR component (i.e. $|\alpha_j| = 1$), leading to random walk-like behaviour.

In this case, the last of these seems most likely, so we attempt to remove it using first differencing.

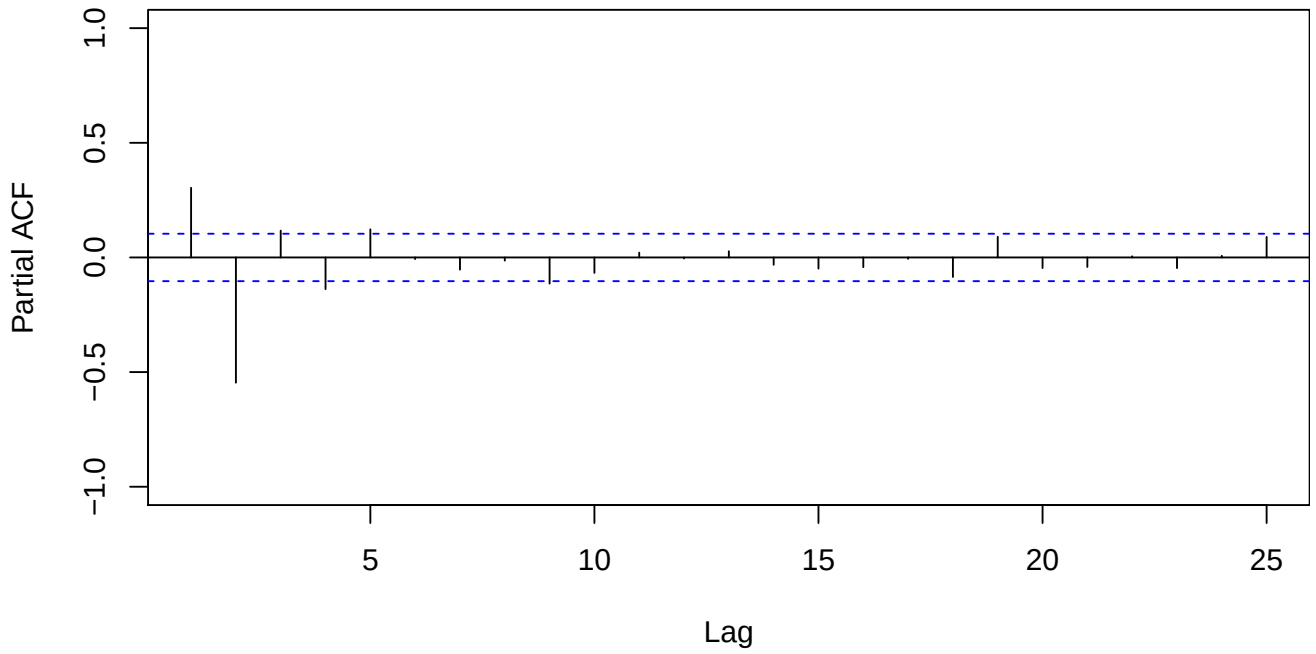
```
X<-diff(Y)
par(mar=c(4,4,1,0))
plot(X,pch=19,cex=0.5);lines(X);abline(h=0,lty=2)
```



```r
acf(X,ylim=range(-1,1))    #ACF
```




```r
pacf(X,ylim=range(-1,1))   #PACF
```



These plots indicate that after differencing, the process seems stationary; the acf plot indicates that the acf diminishes to zero rapidly, and the pacf plot indicates that the pacf is only non-zero for a small number of lags. Therefore an AR(p) model (possibly an ARMA(p, q) model) with p (and q) small should be adequate.

We can then use model selection using AIC or BIC to assess which exact model might be best. Recall that the AIC and BIC calculations essentially assume that the data are Gaussian, so that the log-likelihood for data $\mathbf{x}_{1:n} = (x_1, \ldots, x_n)^\top$ is

$$\ell_n(\theta, \phi, \sigma_Z) = \text{const.} - \frac{1}{2} \log |\Gamma_X(n; \theta, \phi, \sigma_Z)| - \frac{1}{2} \mathbf{x}_{1:n}^\top \{\Gamma_X(n; \theta, \phi, \sigma_Z)\}^{-1} \mathbf{x}_{1:n}.$$

We then perform estimation by maximizing this function with respect to the parameters to obtain estimates $(\widehat{\theta}, \widehat{\phi}, \widehat{\sigma}_Z)$. The AIC and BIC are then defined by

- AIC:
$$\text{AIC}_n(p,q) = -2\ell_n(\widehat{\theta}, \widehat{\phi}, \widehat{\sigma}_Z) + 2(p+q+1)$$
- BIC:
$$\text{BIC}_n(p,q) = -2\ell_n(\widehat{\theta}, \widehat{\phi}, \widehat{\sigma}_Z) + (\log n)(p+q+1)$$

We select the model for which the criterion is **minimized**. First we compare AR(p) models for $p = 0, 1, 2, \ldots, 10$:

```
fit.0<-ar(X,order.max=10,demean=T,method='mle')
fit.0

+
+ Call:
+ ar(x = X, order.max = 10, method = "mle", demean = T)
+
+ Coefficients:
+       1        2        3        4        5
+  0.5669  -0.7076   0.2737  -0.2059   0.1239
+
+ Order selected 5  sigma^2 estimated as  0.4376

fit.0$aic

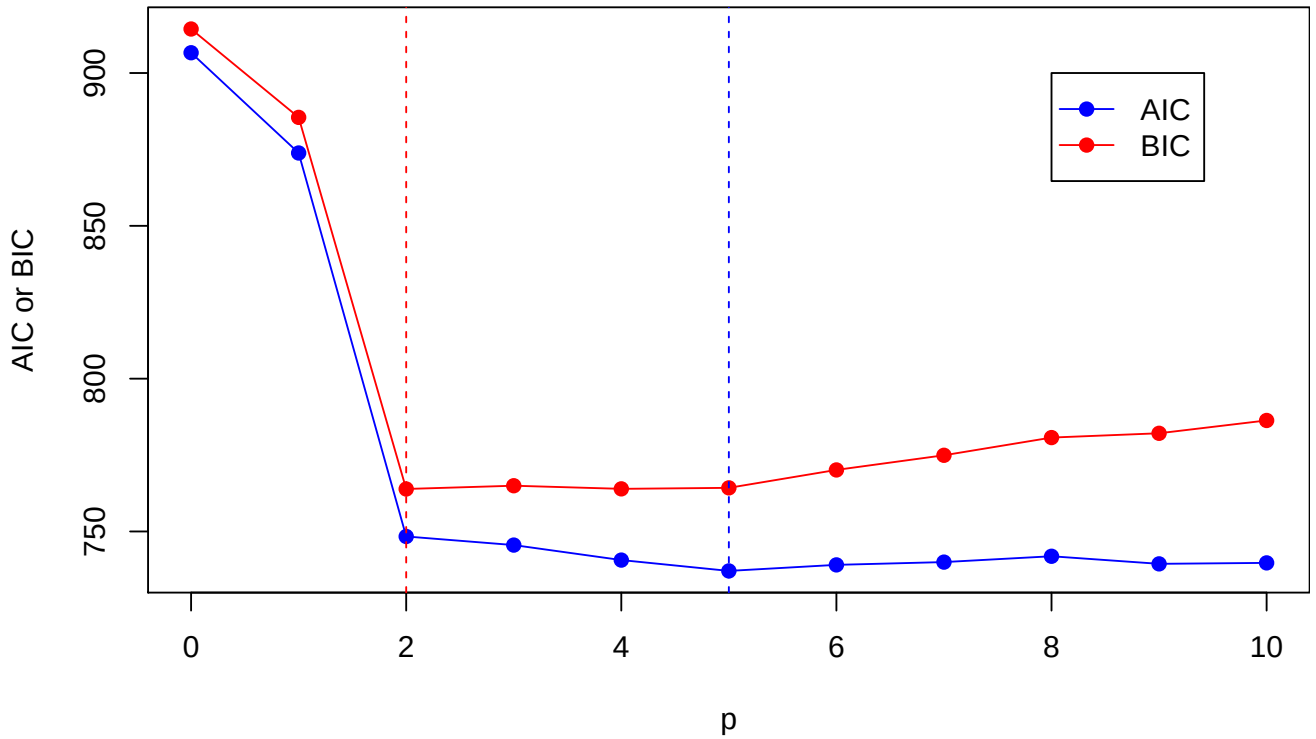
+          0          1          2          3          4          5
+ 169.545130 136.750339  11.275311   8.460292   3.554160   0.000000
+          6          7          8          9         10
+   1.980997   2.866015   4.794448   2.316569   2.721476
```

The AIC values reported here are the values above the minimizing value. These results indicate that an AR(5) model is selected. We can attempt the same analysis using the `arima` function

```
AIC.vec<-BIC.vec<-rep(0,11)
for(p in 0:10){
    fit.p<-arima(X,order=c(p,0,0),include.mean=T,method='ML')
    AIC.vec[1+p]<-fit.p$aic
    BIC.vec[1+p]<-BIC(fit.p)
}
cbind(0:p,AIC.vec,BIC.vec)

+           AIC.vec  BIC.vec
+  [1,]  0 906.6409 914.4075
+  [2,]  1 873.8456 885.4956
+  [3,]  2 748.3710 763.9043
+  [4,]  3 745.5559 764.9725
+  [5,]  4 740.6499 763.9498
+  [6,]  5 737.0955 764.2787
+  [7,]  6 739.0773 770.1438
+  [8,]  7 739.9830 774.9329
+  [9,]  8 741.8955 780.7287
+ [10,]  9 739.4123 782.1288
+ [11,] 10 739.7279 786.3278

par(mar=c(4,4,1,0))
plot(0:p,AIC.vec,pch=19,col='blue',ylim=range(c(AIC.vec,BIC.vec)),xlab='p',ylab='AIC or BIC')
lines(0:p,AIC.vec,col='blue')
points(0:p,BIC.vec,pch=19,col='red');lines(0:p,BIC.vec,col='red')
legend(8,900,c("AIC","BIC"),pch=19,lty=1,col=c('blue','red'))
abline(v=which.min(AIC.vec)-1,col='blue',lty=2)
abline(v=which.min(BIC.vec)-1,col='red',lty=2)
```



Here the AIC and BIC values are on an absolute scale (that is, not relative to the minimizing value). The AIC is minimized at $p = 5$ as before, but the BIC is minimized at $p = 2$. Overall, there seems to be little to choose between the AR(p) models for $p = 2, 3, 4, 5$.

We can also perform the calculation for ARMA(p, q) models.

```
AIC.mat<-BIC.mat<-matrix(0,7,7)
for(p in 0:6){
    for(q in 0:6){
        fit.pq<-arima(X,order=c(p,0,q),include.mean=T,method='ML',optim.control = list(maxit=1000))
        AIC.mat[1+p,1+q]<-fit.pq$aic
        BIC.mat[1+p,1+q]<-BIC(fit.pq)
    }
}

+ Warning in log(s2):  NaNs produced

rownames(AIC.mat)<-paste("p=",c(0:6),sep="");colnames(AIC.mat)<-paste("q=",c(0:6),sep="")
rownames(BIC.mat)<-paste("p=",c(0:6),sep="");colnames(BIC.mat)<-paste("q=",c(0:6),sep="")
```

Again, we can inspect the AIC and BIC values relative to the minimum value:

```
round(AIC.mat-min(AIC.mat),4)

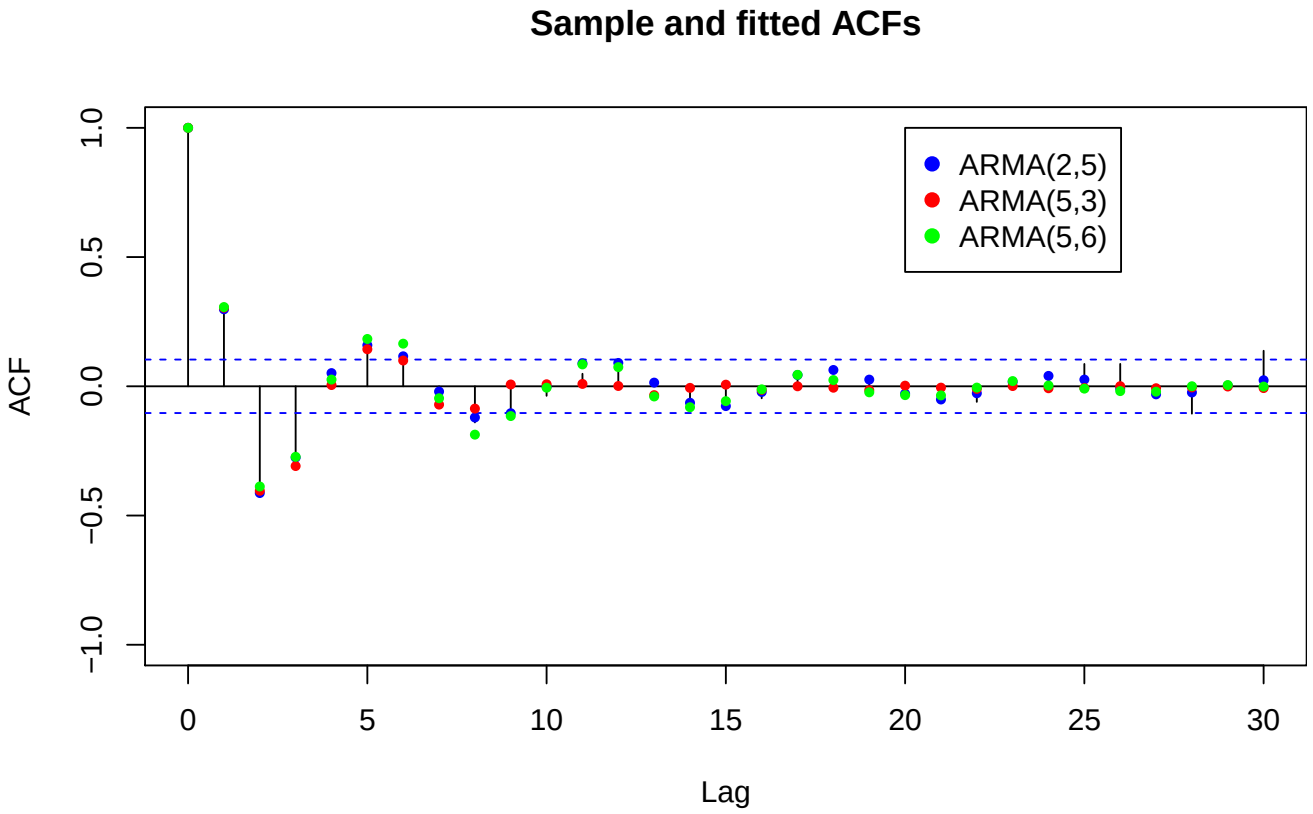
+           q=0     q=1     q=2    q=3    q=4    q=5    q=6
+ p=0 170.8550 48.0151 42.2155 3.1097 5.0853 5.9643 7.5072
+ p=1 138.0598 48.1059 21.6006 5.0921 6.9970 7.8004 9.5016
+ p=2  12.5852  4.7890  2.6154 3.9353 5.1322 0.0000 0.8700
+ p=3   9.7700  3.7525  4.5639 6.6151 5.2325 0.6557 1.3883
+ p=4   4.8640  3.3852  0.9126 1.9124 4.7766 1.6852 2.5864
+ p=5   1.3096  3.3000  4.9654 0.1445 0.2931 1.1030 0.1080
+ p=6   3.2914  5.0534  3.5080 0.6336 1.2691 4.2875 1.9387
```

For the AIC there are several competing models: the minimum is obtained for the ARMA$(2, 5)$ model, however, there are seven other models within 1.00 of this minimizing value. We can compare the fitted acfs of the following models

```r
fit.25<-arima(X,order=c(2,0,5),include.mean=T,method='ML',optim.control = list(maxit=1000))
fit.53<-arima(X,order=c(5,0,3),include.mean=T,method='ML',optim.control = list(maxit=1000))
fit.56<-arima(X,order=c(5,0,6),include.mean=T,method='ML',optim.control = list(maxit=1000))
AIC(fit.25,fit.53,fit.56)

+        df      AIC
+ fit.25  9 735.7859
+ fit.53 10 735.9303
+ fit.56 13 735.8939

acf.25<-ARMAacf(ar=coef(fit.25)[1:2],ma=coef(fit.25)[3:7],lag.max=30)
acf.53<-ARMAacf(ar=coef(fit.53)[1:5],ma=coef(fit.53)[6:8],lag.max=30)
acf.56<-ARMAacf(ar=coef(fit.56)[1:5],ma=coef(fit.56)[6:11],lag.max=30)
par(mar=c(4,4,3,0))
acf(X,lag.max=30,ylim=range(-1,1),main='Sample and fitted ACFs')
points(0:30,acf.25,pch=19,cex=0.6,col='blue')
points(0:30,acf.53,pch=19,cex=0.6,col='red')
points(0:30,acf.56,pch=19,cex=0.6,col='green')
legend(20,1,c('ARMA(2,5)','ARMA(5,3)','ARMA(5,6)'),col=c('blue','red','green'),pch=19)
```



The fitted ACF values are very similar, and match the sample acf values reasonably well.

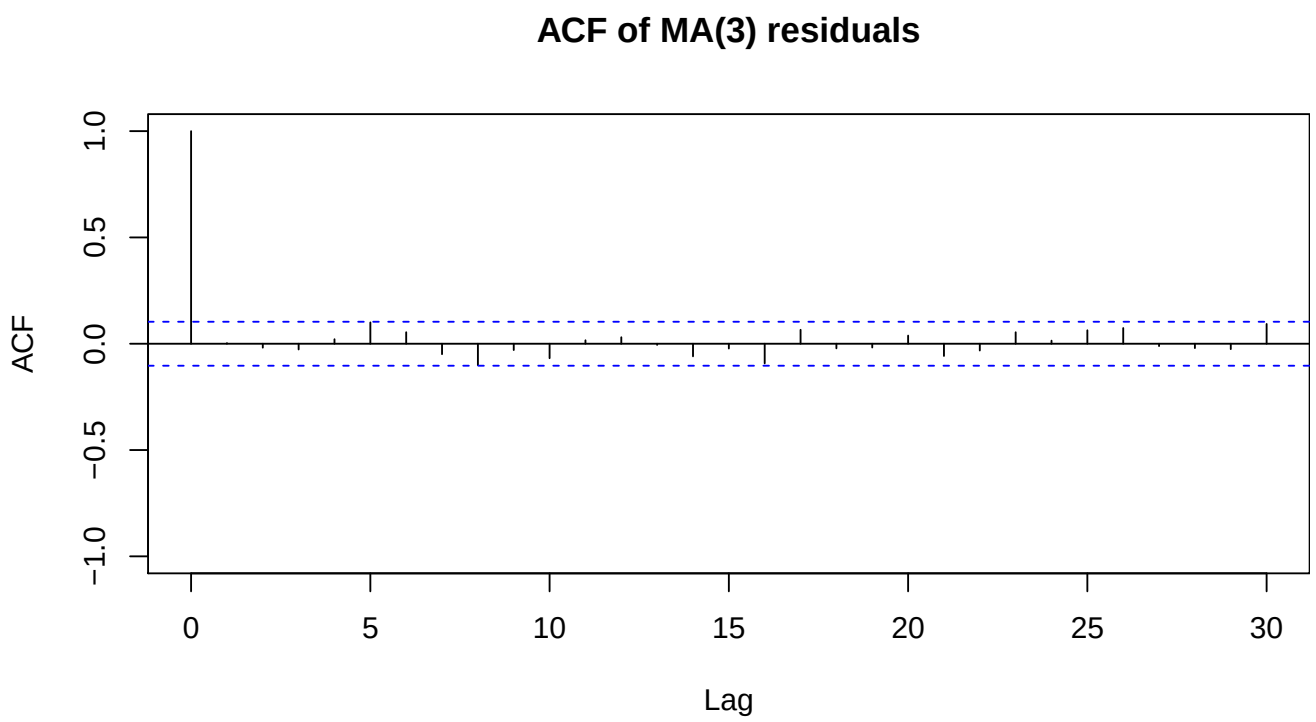
For the BIC, the selected model is the MA(3), and only the ARMA$(2,1)$ model is close in BIC terms.

```r
round(BIC.mat-min(BIC.mat),4)

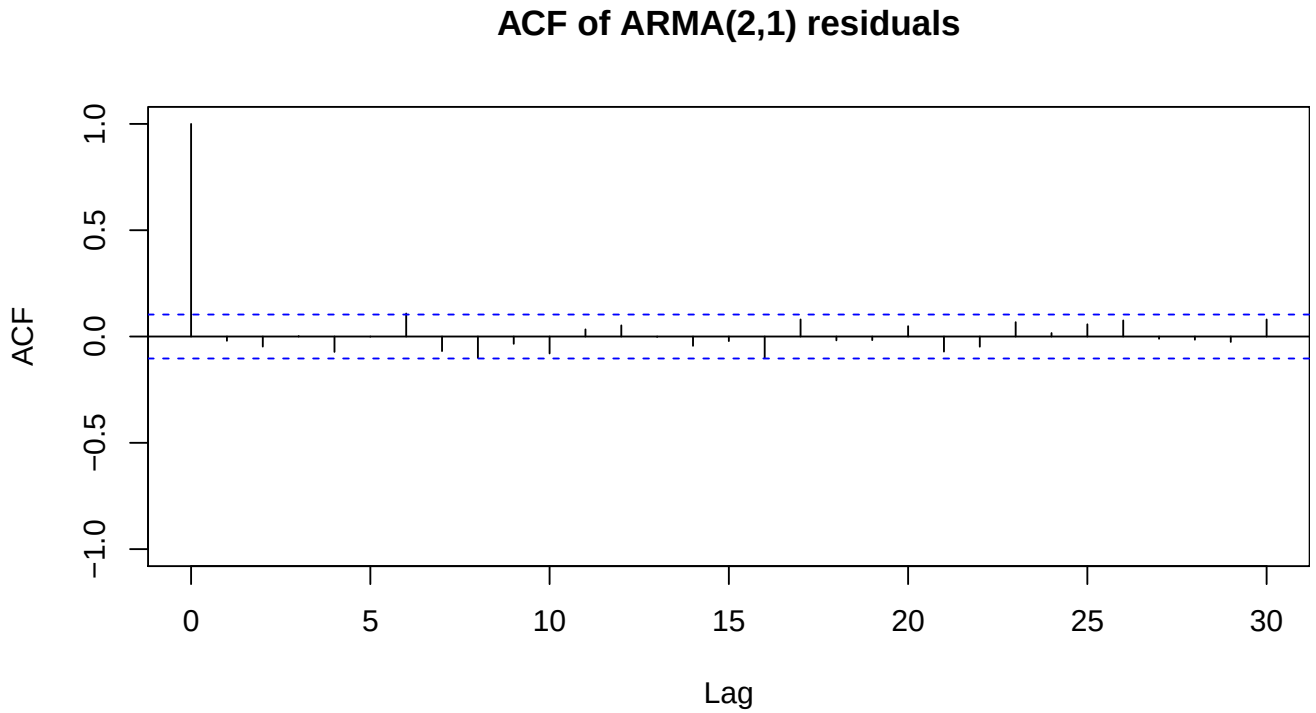
+           q=0      q=1      q=2      q=3      q=4      q=5      q=6
+ p=0 156.0953 37.1388 35.2225  0.0000  5.8590 10.6213 16.0475
+ p=1 127.1834 41.1129 18.4909  5.8657 11.6540 16.3406 21.9252
+ p=2   5.5922  1.6793  3.3890  8.5923 13.6725 12.4236 17.1769
+ p=3   6.6603  4.5261  9.2208 15.1554 17.6561 16.9626 21.5785
+ p=4   5.6377  8.0422  9.4528 14.3360 21.0835 21.8754 26.6600
+ p=5   5.9666 11.8402 17.3890 16.4514 20.4833 25.1766 28.0649
+ p=6  11.8317 17.4770 19.8149 20.8238 25.3427 32.2444 33.7789
```

We may compute the residual series from the fits of these models and compare them

```
fit.03<-arima(X,order=c(0,0,3),include.mean=T,method='ML',optim.control = list(maxit=1000))
fit.21<-arima(X,order=c(2,0,1),include.mean=T,method='ML',optim.control = list(maxit=1000))
res.03<-residuals(fit.03);res.21<-residuals(fit.21)
par(mar=c(4,4,3,0))
acf(res.03,lag.max=30,ylim=range(-1,1),main='ACF of MA(3) residuals')
```



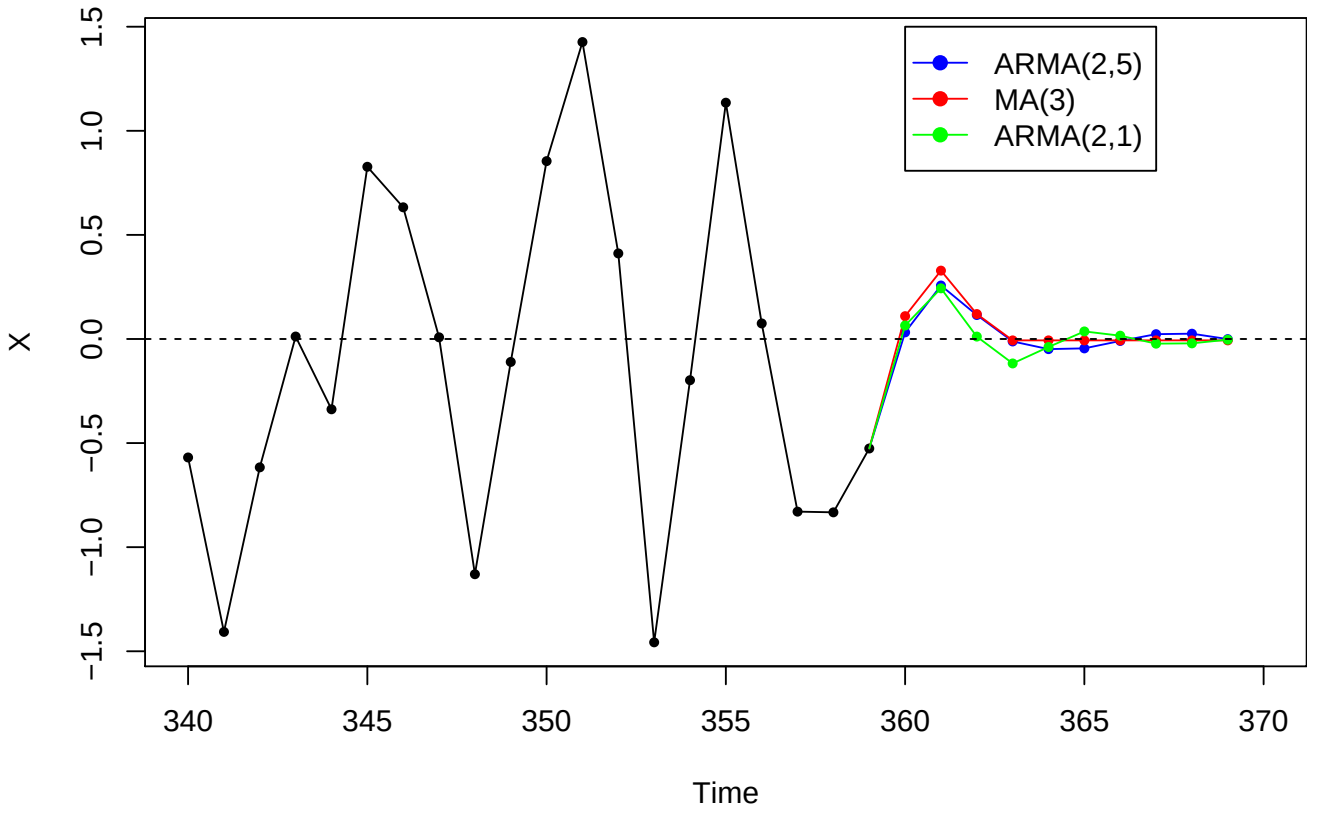

```
acf(res.21,lag.max=30,ylim=range(-1,1),main='ACF of ARMA(2,1) residuals')
```



Both of the residual series seem to be plausibly white noise processes.

Finally, we might compare forecasts from several of these models: for this, we can use the `predict` function:

```
pred.25<-predict(fit.25,n.ahead=10)$pred
pred.03<-predict(fit.03,n.ahead=10)$pred
pred.21<-predict(fit.21,n.ahead=10)$pred
par(mar=c(6,4,1,0))
plot(340:359,X[340:359],xlim=range(340,370),pch=19,cex=0.6,xlab='Time',ylab='X')
lines(340:359,X[340:359])
points(360:369,pred.25,col='blue',pch=19,cex=0.6);lines(359:369,c(X[359],pred.25),col='blue')
points(360:369,pred.03,col='red',pch=19,cex=0.6);lines(359:369,c(X[359],pred.03),col='red')
points(360:369,pred.21,col='green',pch=19,cex=0.6);lines(359:369,c(X[359],pred.21),col='green')
abline(h=0,lty=2)
legend(360,1.5,c('ARMA(2,5)','MA(3)','ARMA(2,1)'),col=c('blue','red','green'),pch=19,lty=1)
```



The forecasts are very similar overall due to the fact that the autocovariances for the different models are similar. As usual, the forecasts return to the process mean of zero as h increases.