CSC 425 Time Series Analysis HW2

Akhil Kumar Ramasagaram

Tuesday, September 29, 2015

Problem 1

Consider the following AR(2) time series process:$r_t = 0.5 + .16r_{t-2} + a_t$, where$\{a_t\}$is a Gaussian white noise series with mean zero and constant variance $\sigma^2 = 0.02$. Note that the AR(2) process has coefficient $\phi_1 = 0$ for r_{t-1}.

a) What is the mean of the time series r_t.

An AR(2) models assumes the form $x_t = \phi_o + \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$. From AR(1), we can assume that mean is,

$$E(x_t) = \mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}$$

$$\mu = \frac{0.5}{1 - 0 - 0.16} = \frac{0.5}{0.84}$$

$$\mu = 0.5952$$

b) Determine if the AR(2) model is stationary.

```
library(zoo)
library(fBasics)
library(lmtest)
library(forecast)
orange = read.table("oranges.csv",header=T, sep=',')
m2=arima(orange$price, order=c(2,0,0))
coeftest(m2)
```

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value  Pr(>|z|)
## ar1          1.066436   0.086679 12.3032 < 2.2e-16 ***
## ar2         -0.278314   0.087693 -3.1737  0.001505 **
## intercept 898.085723  38.497313 23.3285 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
p = polyroot(c(1, -m2$coef[1:3]))
Mod(p)
```

```
## [1] 0.09993033 0.10555831 0.10555831
```

If we have a look at take the absolute values of these polynomials, the weights should be less than 1 to say that the AR model is stationary. Our absolute values are 0.0999303, 0.1055583, 0.1055583. Since the values are less than 1, we can say that the AR(2) model is stationary.

c) Assume that $r_{100} = -0.01$ and $r_{99} = 0.02$. Compute the 1-step and 2-step ahead forecasts of the AR(2) series at the forecast origin $t = 100$.

The 1-step ahead forecast will be, $r_101 = 0.5 + 0.16 * 0.02 = 0.5032$. The 2-step ahead forecast will be, $r_102 = 0.5 + 0.16 * (0.01) = 0.4984$.

d) Compute the lag-1 and lag-2 autocorrelations of r_t.

Lag-1 can be computed using.

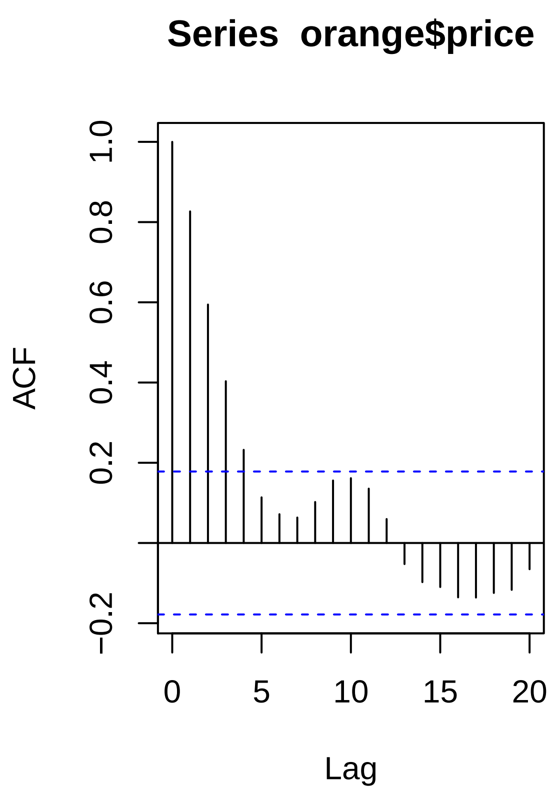
$$\rho_1 = \Phi_1/(1 - \Phi_2)$$

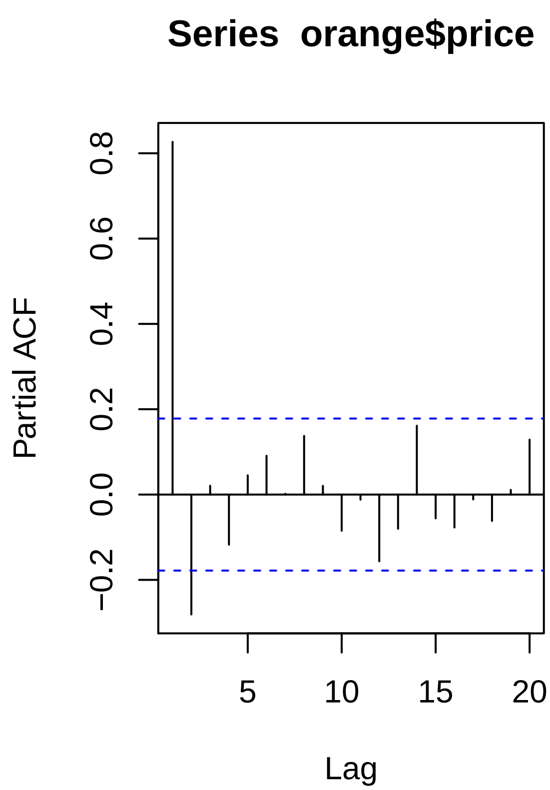
Lag-2 can be computed using.

$$\rho_2 = \rho_1\Phi_1 + \rho_0\Phi_2$$

##Problem 2 ###a) Analyze the ACF and the PACF plots for the orange sales data and discuss which order p model is suggested by the plots?

```
par(mfrow = c(1,2))
acf(orange$price)
pacf(orange$price)
```





From the acf plot I would say that we should use the second order p value.

b) Fit an adequate AR model:

- Examine the significance of the model coefficients, and discuss if all coefficients are significantly different from zero.

Let's build an AR(1) model.

```
m1=Arima(orange$price,order=c(1,0,0))
coeftest(m1)
```

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value  Pr(>|z|)
## ar1         0.836211   0.049304  16.960 < 2.2e-16 ***
## intercept 890.311902  50.859893  17.505 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

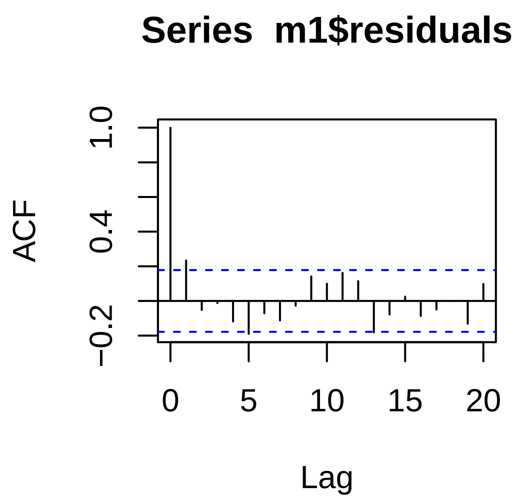
```
ar(orange$price)
```

```
##
## Call:
## ar(x = orange$price)
##
## Coefficients:
##      1       2
##  1.059  -0.281
##
## Order selected 2  sigma^2 estimated as  8911
```

The first coefficient is not zero and the second coefficient is closer to zero.

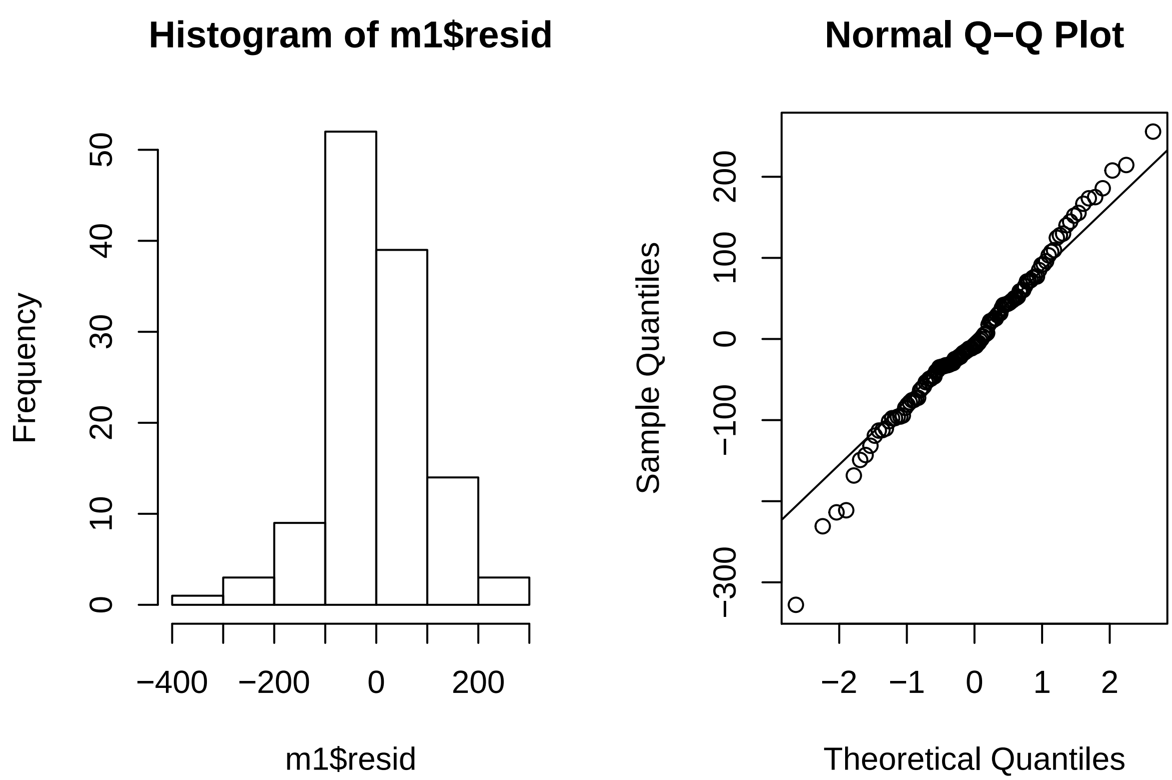
- Perform a residual analysis and discuss if the selected model is adequate.
- Compute ACF functions of residuals and test if residuals are white noise.

```
acf(m1$residuals)
```



It looks like there is no white noise in the data. + Plot histogram and normal quantile plots of residuals.

```
par(mfrow = c(1,2))
hist(m1$resid)
qqnorm(m1$resid)
qqline(m1$resid)
```



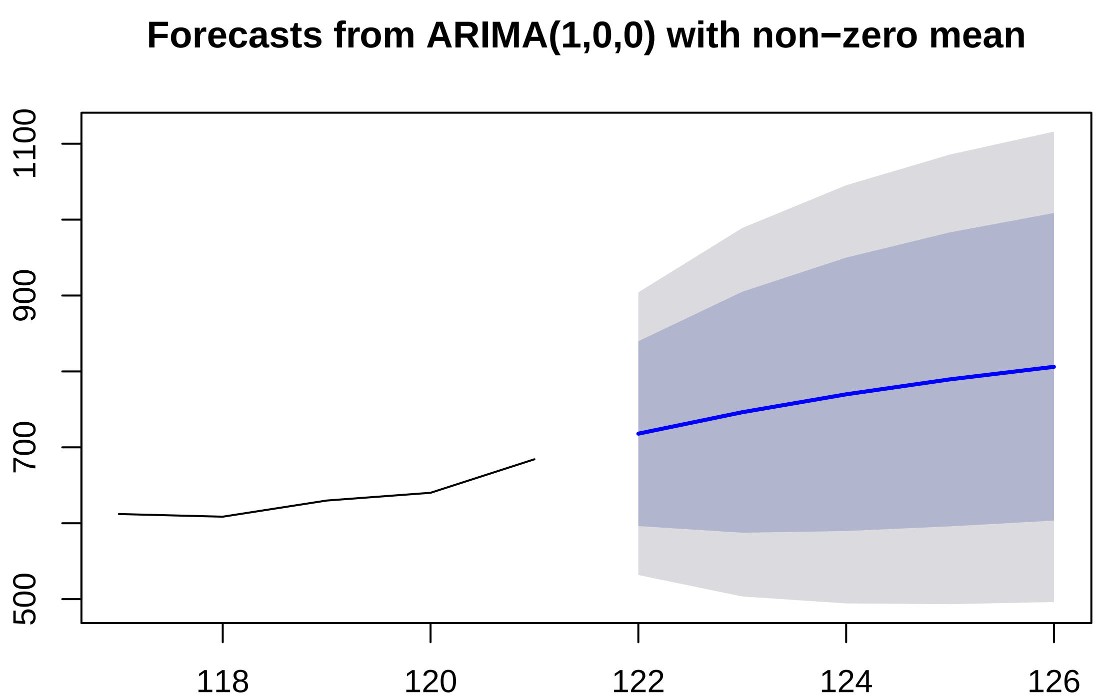
We can say that the variance of residuals are high, the distribution is skewed towards left. The distribution is not normal, and cannot describe the orange sales with this distribution.

c) The Expression for the above AR model can be writtern as.

$$X_t = 145.8381 + 08362X_{t-1} + a_t$$

d) Use the selected AR model to compute up to 5 step-ahead forecasts starting from the last observation in the dataset. Write down the forecasts and their margin of errors or prediction intervals.

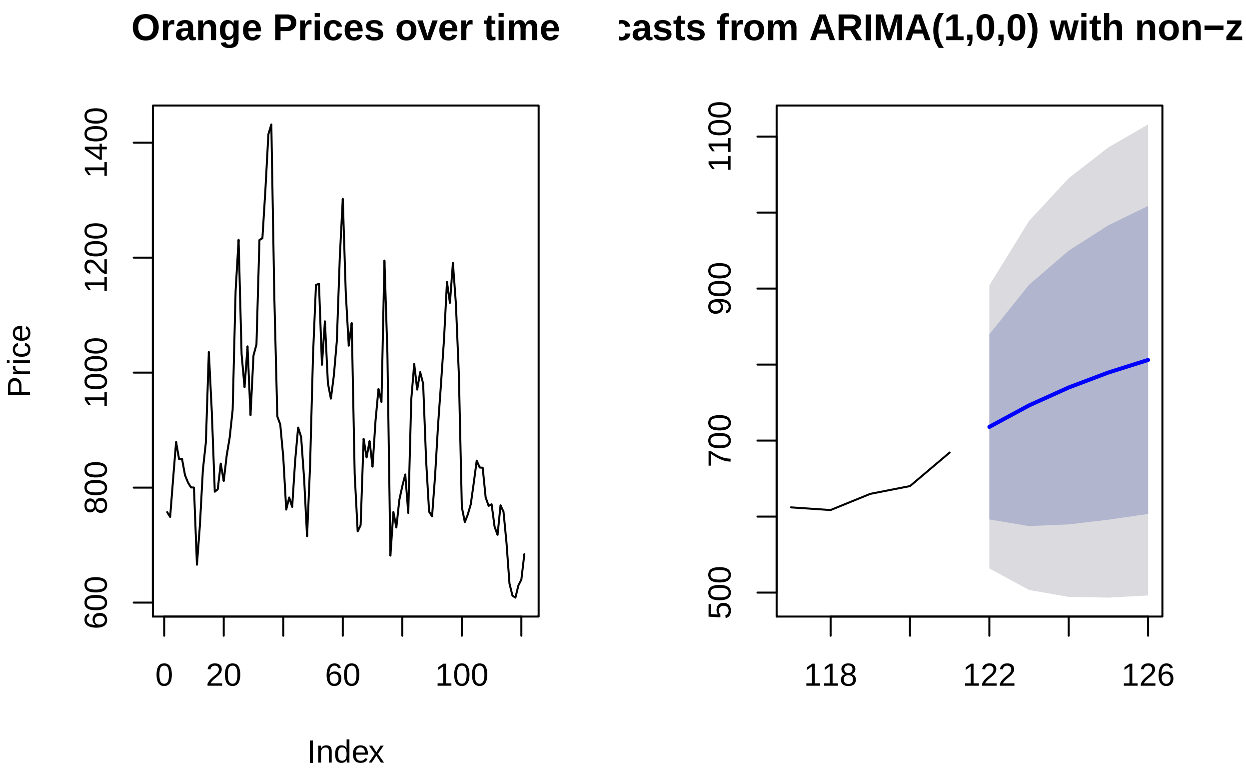
```
f=forecast.Arima(m1,h=5)
plot(f,include=5)
```



Here the forcast is showing an upward trend and also converging to the mean.

e) Discuss where the forecasts converge to after a few steps.

```
par(mfrow = c(1,2))
plot(orange$price, type = "l", ylab = "Price", main = "Orange Prices over time")
f_5=forecast.Arima(m1,h=5)
plot(f_5,include=5)
```



The mean of the orange prices is 904.5679339. From the forecast plot we can that the forecast is slightly increasing towards 900 and then converging around the mean.