
STAT 463: Homework 4

Name:Kyle Salitrik | **ID#:** 997543474 | **PSU ID:** *kps168*

February 22, 2018

PROBLEM 1

When forecasting, for $t > T \Rightarrow e_t = 0$

$$y_{81} = 50 - 0.7 * 3 = 47.9$$

$$y_{82} = 50 - 0.7 * 47.9 = 16.47$$

$$y_{83} = 50 - 0.7 * 16.47 = 38.471$$

a)

For y_{81} 's forecast see above equations.

$$\begin{aligned} \text{PI}y_{81} &= X_{81}^{80} \pm 1.96 * \sqrt{\hat{\sigma}_w^2 * \Psi_0^2} \\ &= 47.9 \pm 1.96 * \sqrt{2.5^2 * 1} \\ &= \{43, 52.8\} \end{aligned}$$

b)

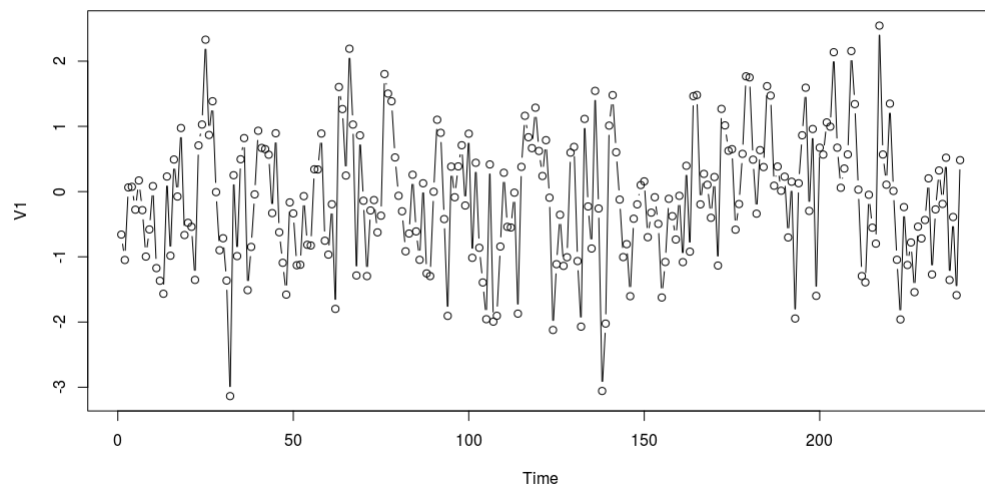
For y_{83} 's forecast see above equations.

$$\begin{aligned} \text{PI}y_{83} &= X_{83}^{80} \pm 1.96 * \sqrt{\hat{\sigma}_w^2 * (\Psi_0^2 + \Psi_1^2 + \Psi_2^2)} \\ &= 38.471 \pm 1.96 * \sqrt{2.5^2 * (1 + 0.49 + 0.2401)} \\ &\approx \{32.03, 44.92\} \end{aligned}$$

PROBLEM 2

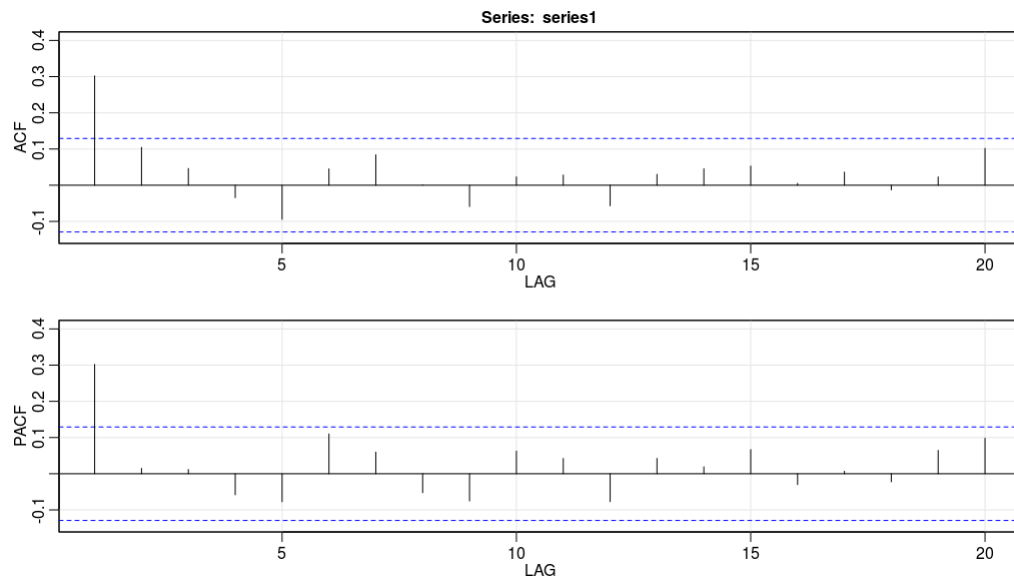
a)

Plotting the time series doesn't appear to show any significant trend. The plot appears to be stationary as well.



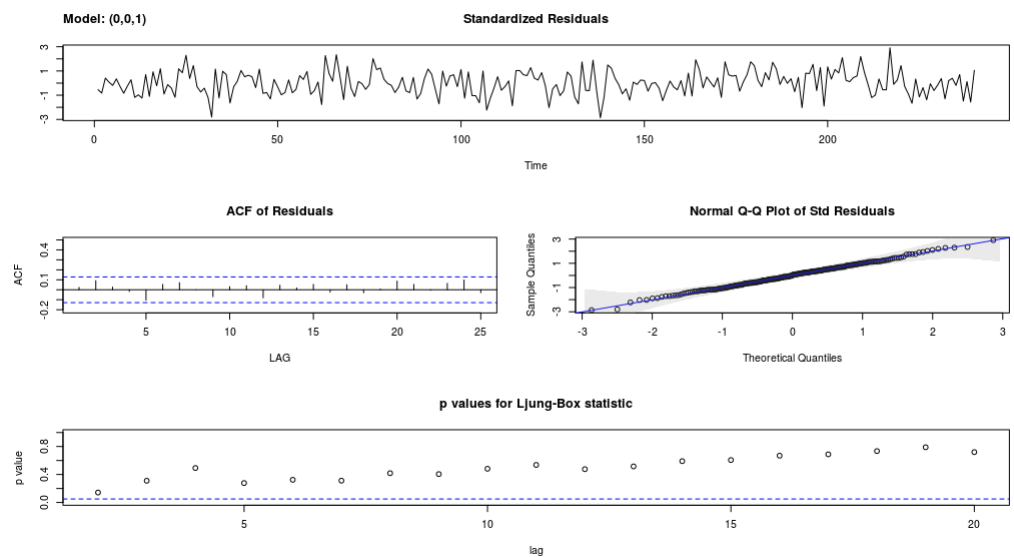
b)

Looking at the ACF and PACF plots, there is a single significant spike in each indicating that the model is possibly AR(1), MA(1) or ARMA(1,1).



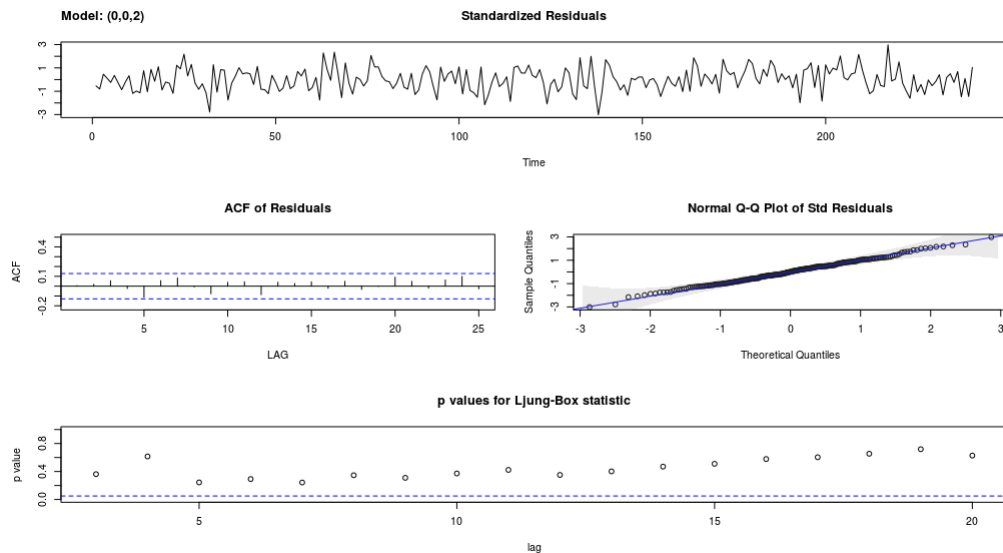
c)

The MA(1) model seems suitable based on the diagnostic plots. The sigma squared value is also 0.9303.



d)

While the MA(2) model may be suitable, it does have a lower sigma squared value of 0.9242.



e)

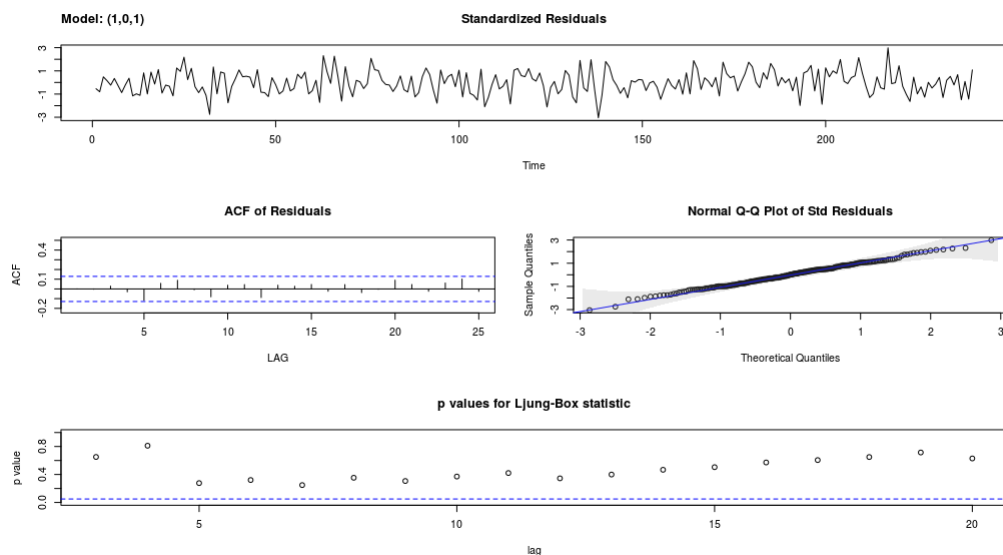
In both models, the MA1 coefficient is significant, having a 0 p-value in both cases. However in the MA2 model, the p-value of the MA2 coefficient is 0.207, which is fairly insignificant, indicating that MA1 is a better option.

f)

$$y_t = -0.1124 + 0.2873y_{t-1} + e_t$$

g)

Using an ARIMA(1,0,1) model results in a sigma squared value of 0.9219, lower than the AR(1) model. However the AR(1) coefficient has a significantly lower p-value than the MA(1): 0.0877 vs 0.8212



h)

Based purely on the sigma-squared values, the AR(1) model is better. In this model the coefficient is also more significant than either coefficient in the ARIMA(1,0,1) model.

i)

```
> sarima.for(series1,6, 0,0,2)
$pred
Time Series:
Start = 241
End = 246
Frequency = 1
[1] 0.07286824 -0.03681730 -0.11285782 -0.11285782 -0.11285782 -0.11285782

$se
Time Series:
Start = 241
End = 246
Frequency = 1
[1] 0.9613769 1.0018247 1.0045077 1.0045077 1.0045077 1.0045077
```

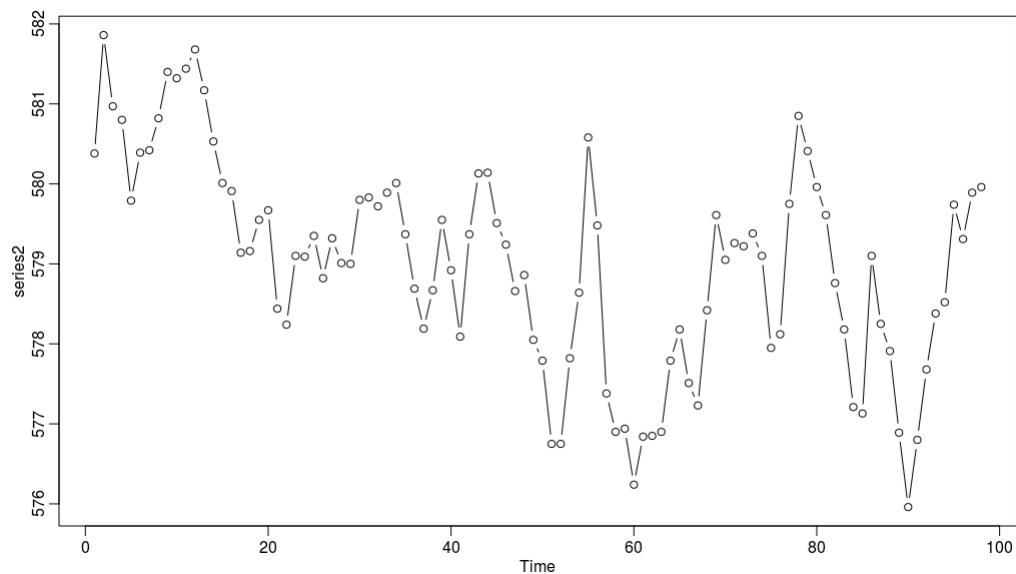
j)

The values converge due to the properties of the time series model itself. At infinite time, the values of the prediction of the time series will converge to the mean of the series.

PROBLEM 3

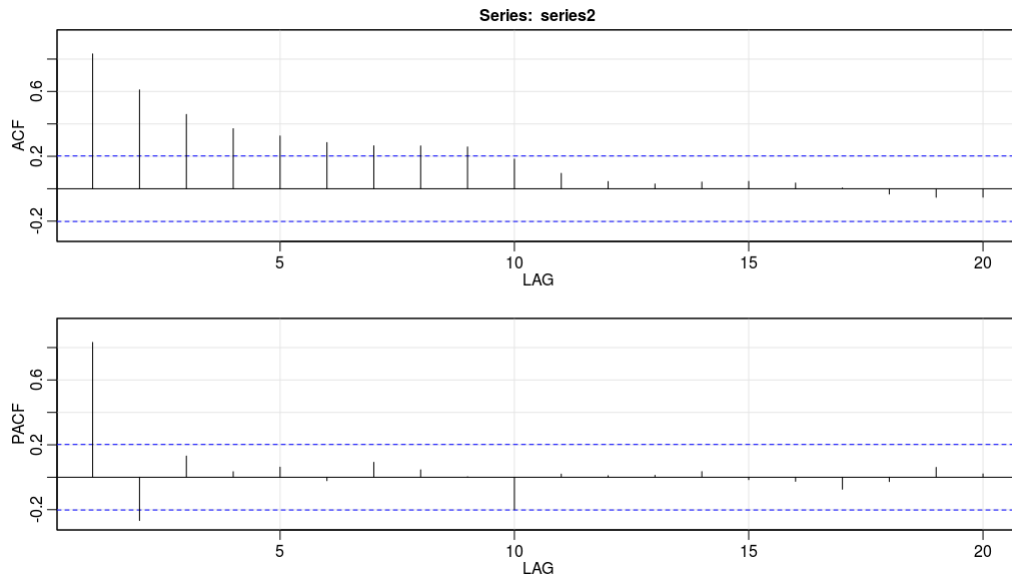
a)

Looking at the time series, there appears to be a seasonality to the data.



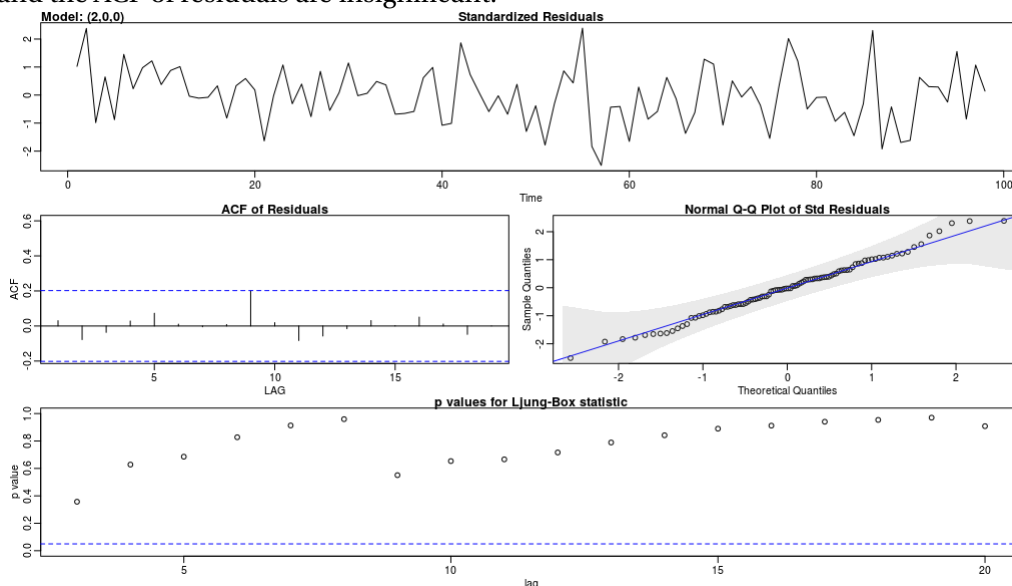
b)

Looking at the ACF and PACF plots, there is a single significant spike in each indicating that the model is possibly AR(2) or anywhere from MA(1) to MA(9).



c)

The AR(2) model appears to be fairly suitable for use as all of the points within the Ljung-Box plot and the ACF of residuals are insignificant.



d)

Based on the results of the sarima function, the AR(1) coefficient has a p-value of 0 and the AR(2) coefficient has a p-value of 0.0151. Depending on the threshold one wants to use, these both may be significant or only the AR(1).

e)

Using the given equation, we can estimate the intercept to be: $\delta = 579.0473(1 - 1.0436 + 0.2495) \approx 119.23$

And the regression equation is as follows: $y_t = 119.23 + 1.0436 * y_{t-1} - 0.2495 * y_{t-2} + e_t$

f)

For the estimates:

$$y_{99} = 119.23 + 1.0436 * 579.89 - 0.2495 * 579.96 = 579.70$$

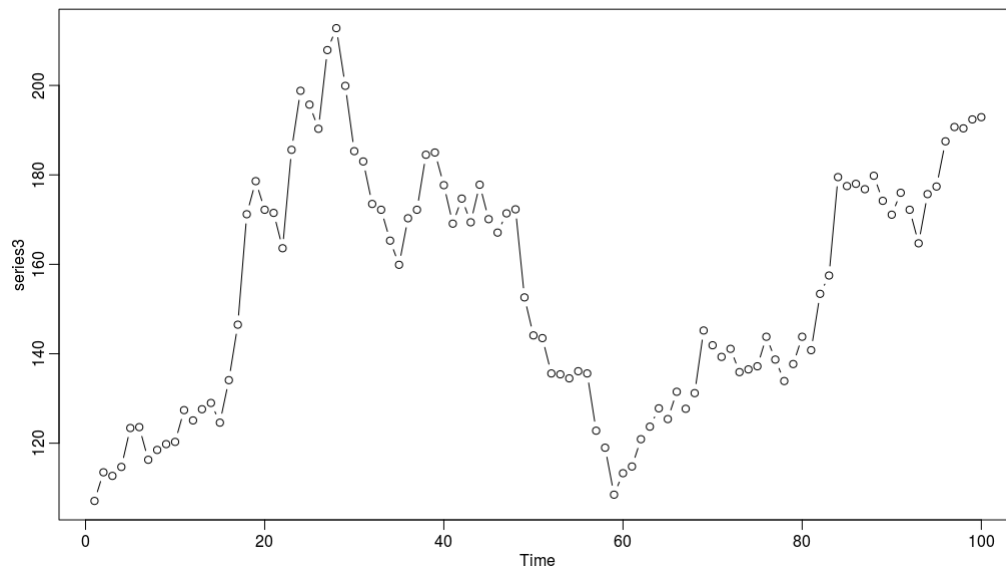
$$y_{100} = 119.23 + 1.0436 * 579.70 - 0.2495 * 579.96 = 579.50$$

$$y_{101} = 119.23 + 1.0436 * 579.50 - 0.2495 * 579.70 = 579.36$$

PROBLEM 4

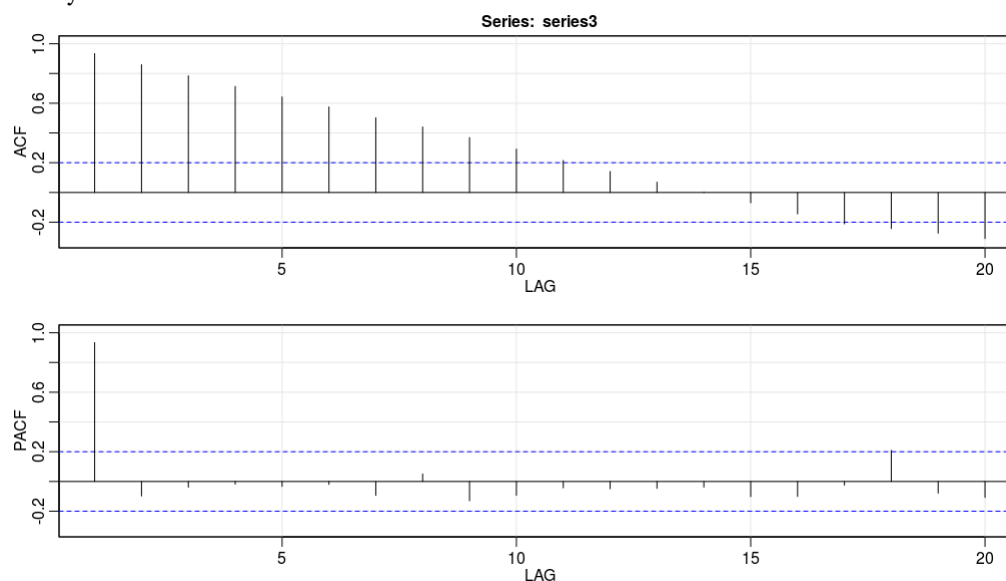
a)

There appears to be a yearly trend for the price of flour, however it seems that each year is mostly independent.



b)

An AR(1) model may apply based on the PACF, however it's more likely that a high order MA model may fit best.



c)

For the difference, there is no significant spikes, meaning that no time series related models may fit the first difference of the data.

