Homework 5

Deanna Springgay

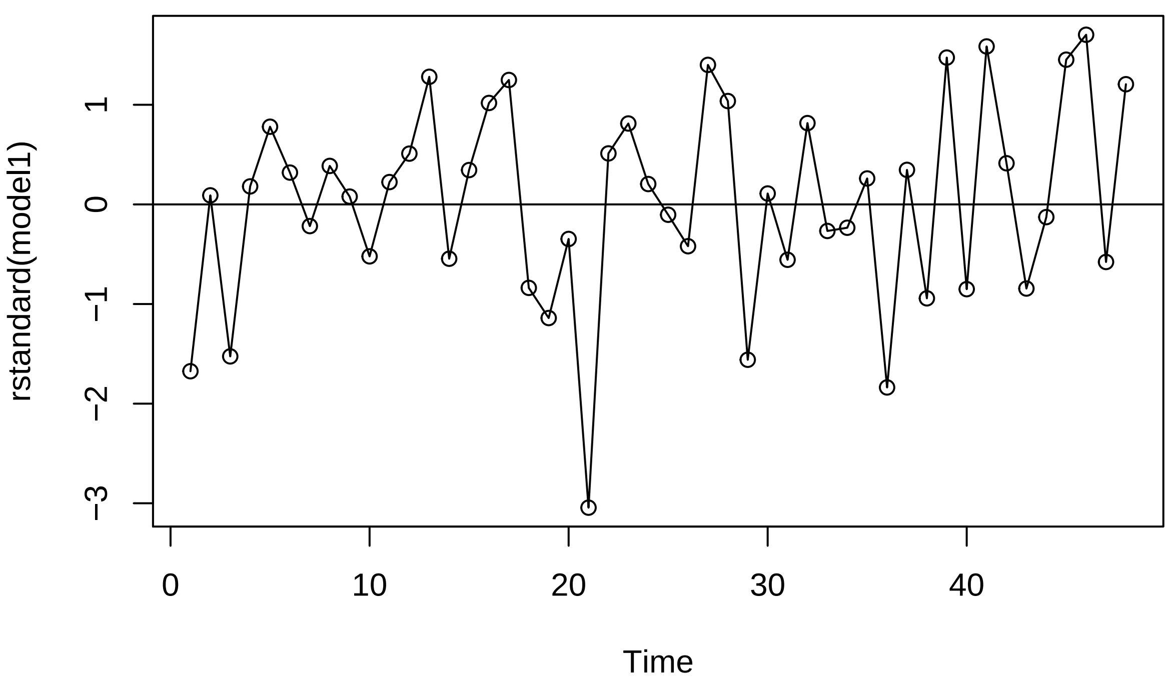
4/22/2021

Problems

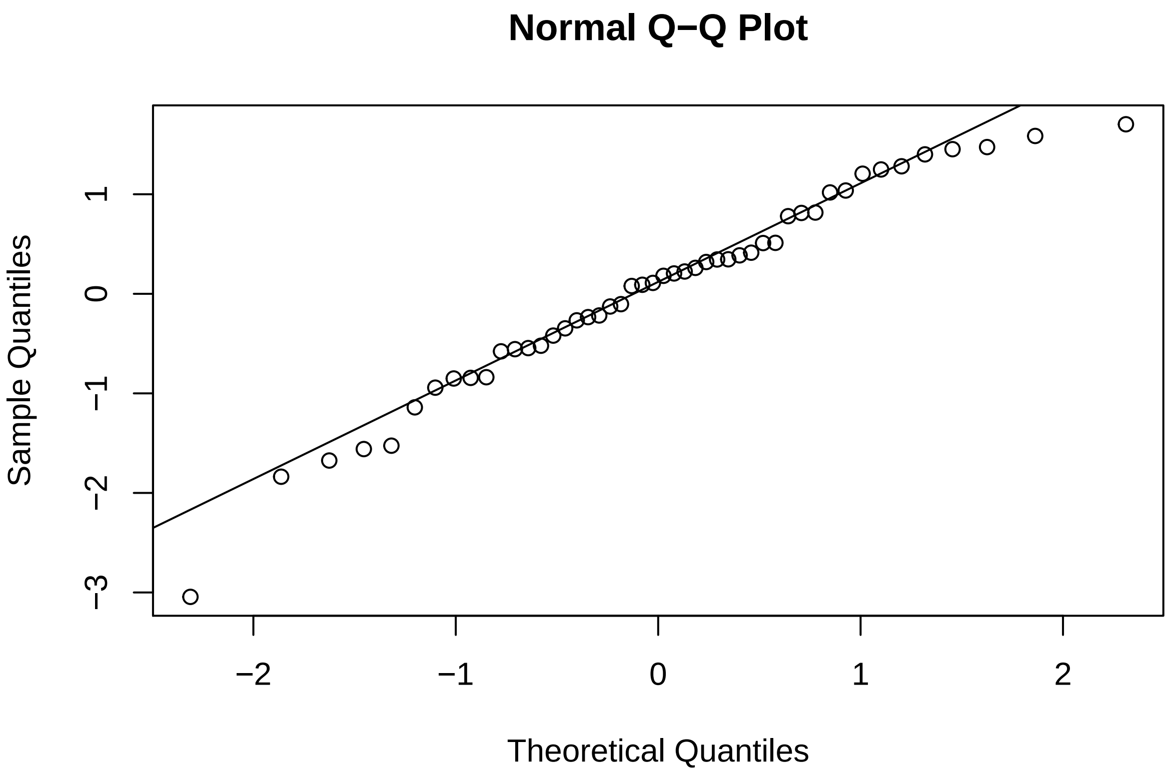
Written questions are attached at the end

8.6

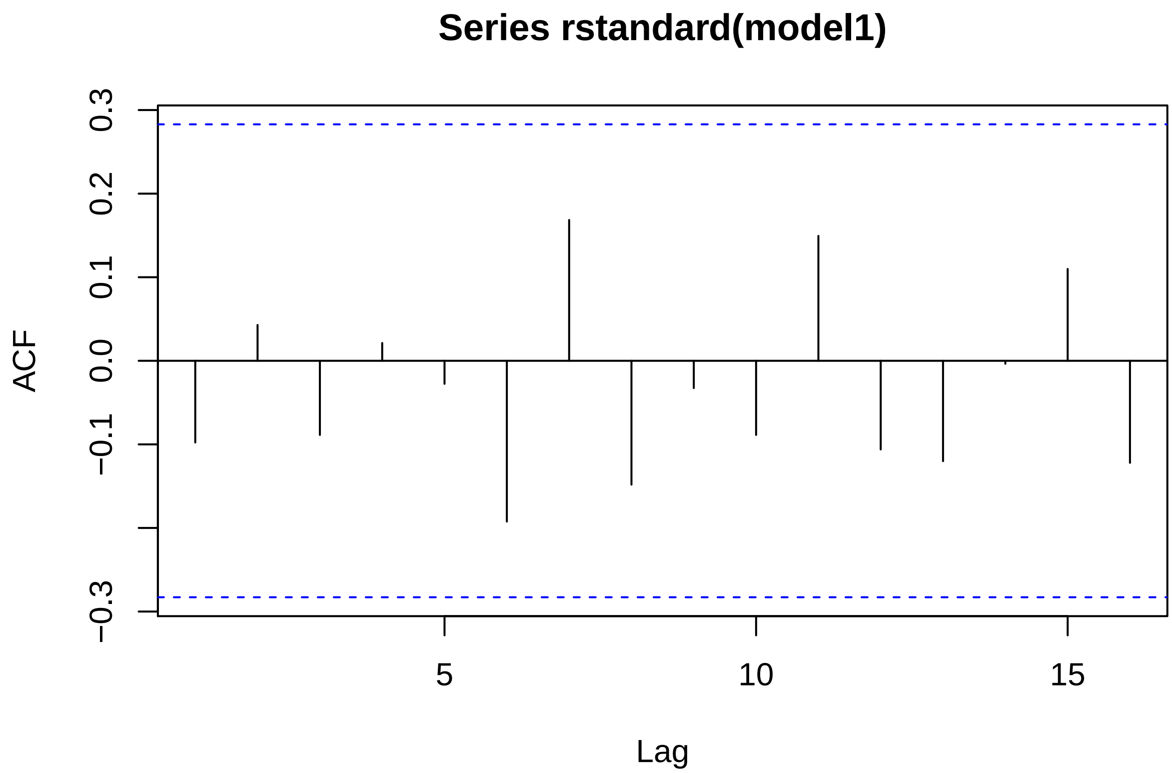
a) The residuals seem to be random around a zero mean.



b) The residuals seem to follow a normal distribution.



c) None of the lags are significant.



d) From the Ljung-Box test we fail to reject the null hypothesis of independent residuals.

```
##
##  Box-Ljung test
##
## data:  residuals from  model1
## X-squared = 8.9256, df = 10, p-value = 0.5392
```

8.9

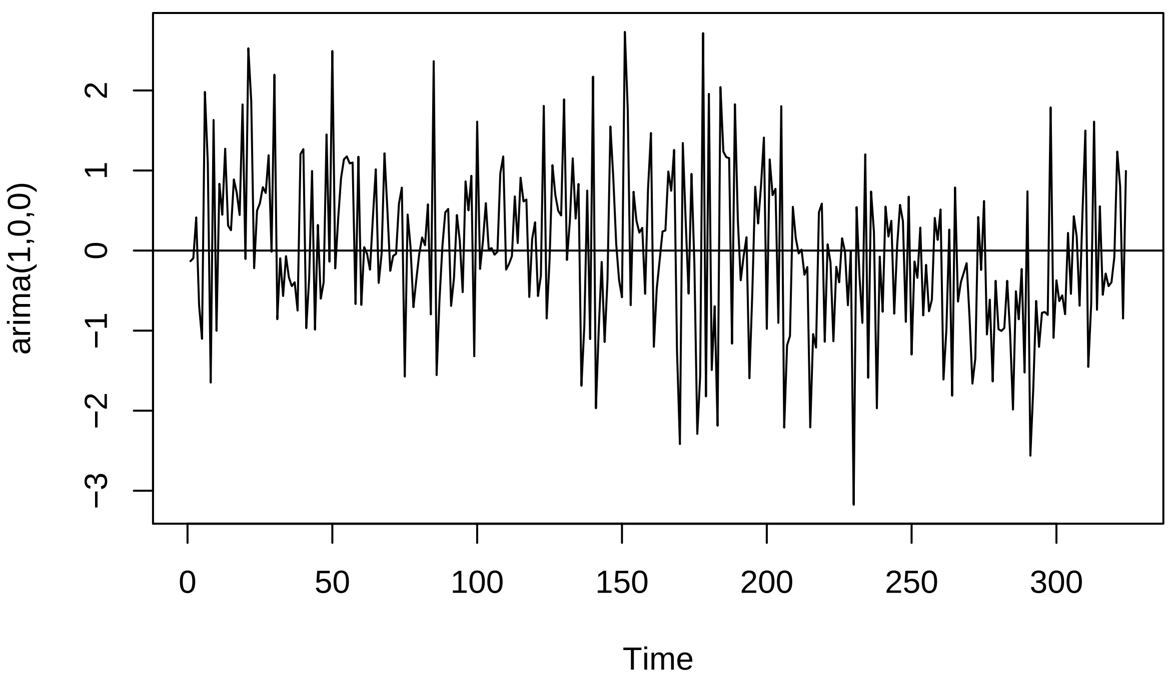
While both models are significant, the arima(0,1,1) model has a better log likelihood and AIC.

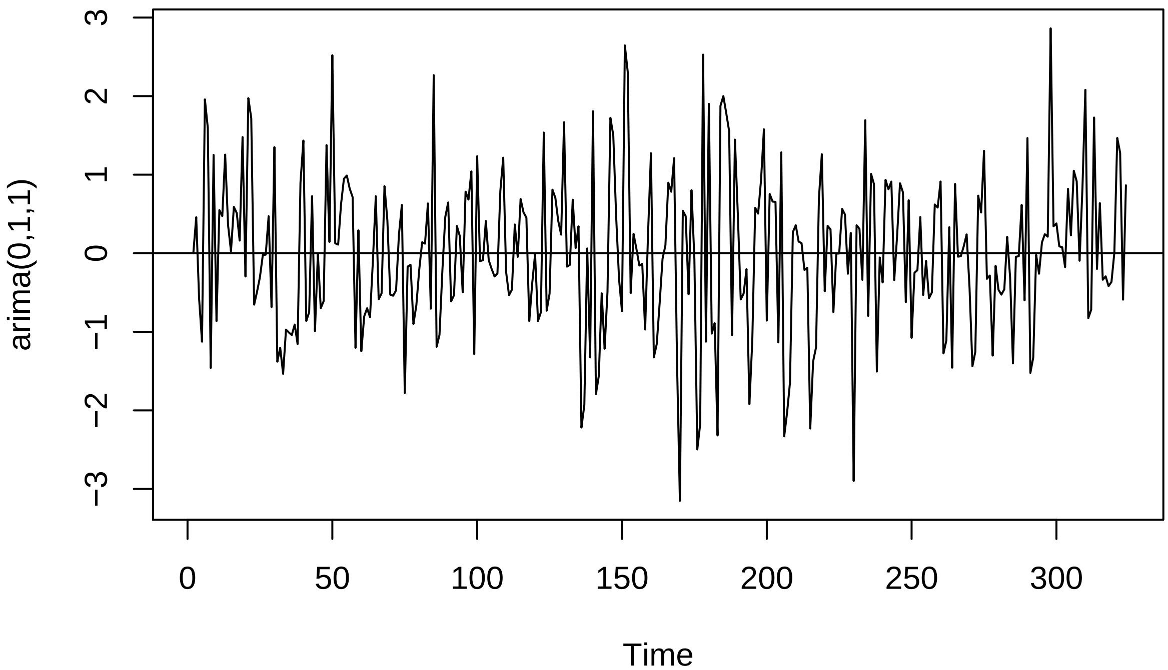
```
##
## Call:
## arima(x = robot, order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.3074     0.0015
## s.e.  0.0528     0.0002
##
## sigma^2 estimated as 6.482e-06:  log likelihood = 1475.54,  aic = -2947.08
```

```
##
## Call:
## arima(x = robot, order = c(0, 1, 1))
##
## Coefficients:
```

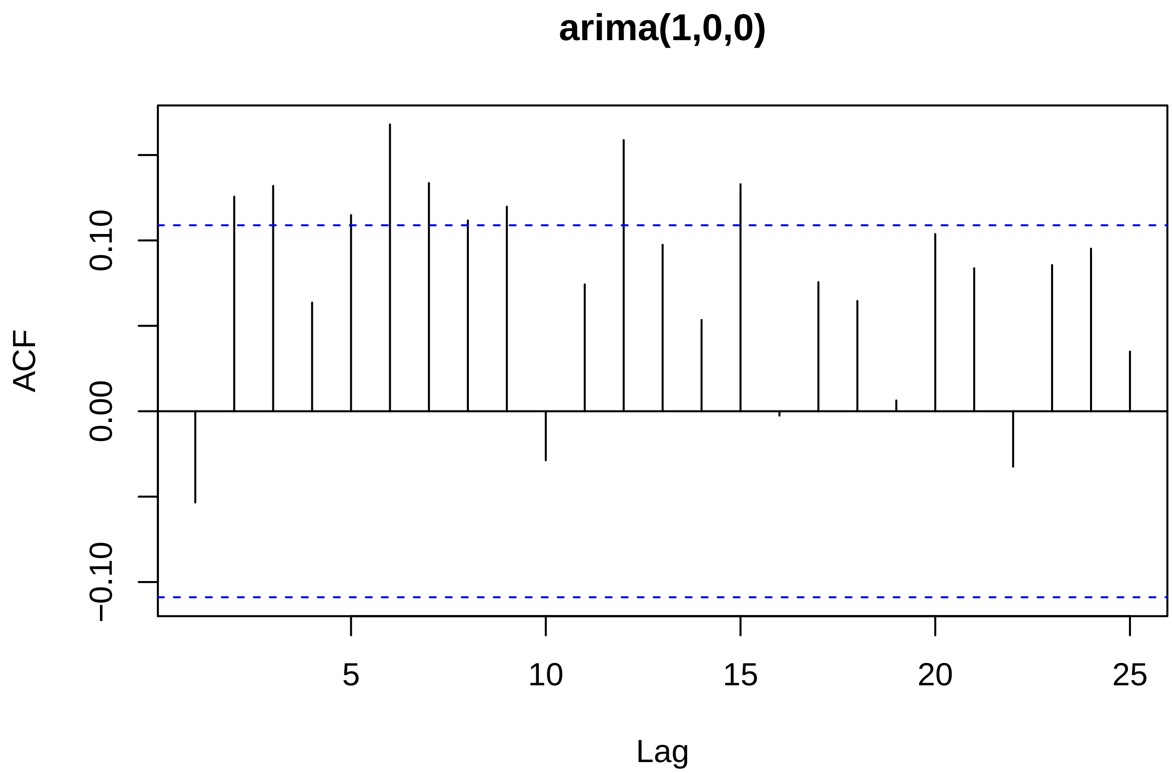
```
##            ma1
##        -0.8713
## s.e.    0.0389
##
## sigma^2 estimated as 6.069e-06:  log likelihood = 1480.95,  aic = -2959.9
```

The arima(1,0,0) model might have some drift, but arima(0,1,1) doesn't seem to have any problems at this point.

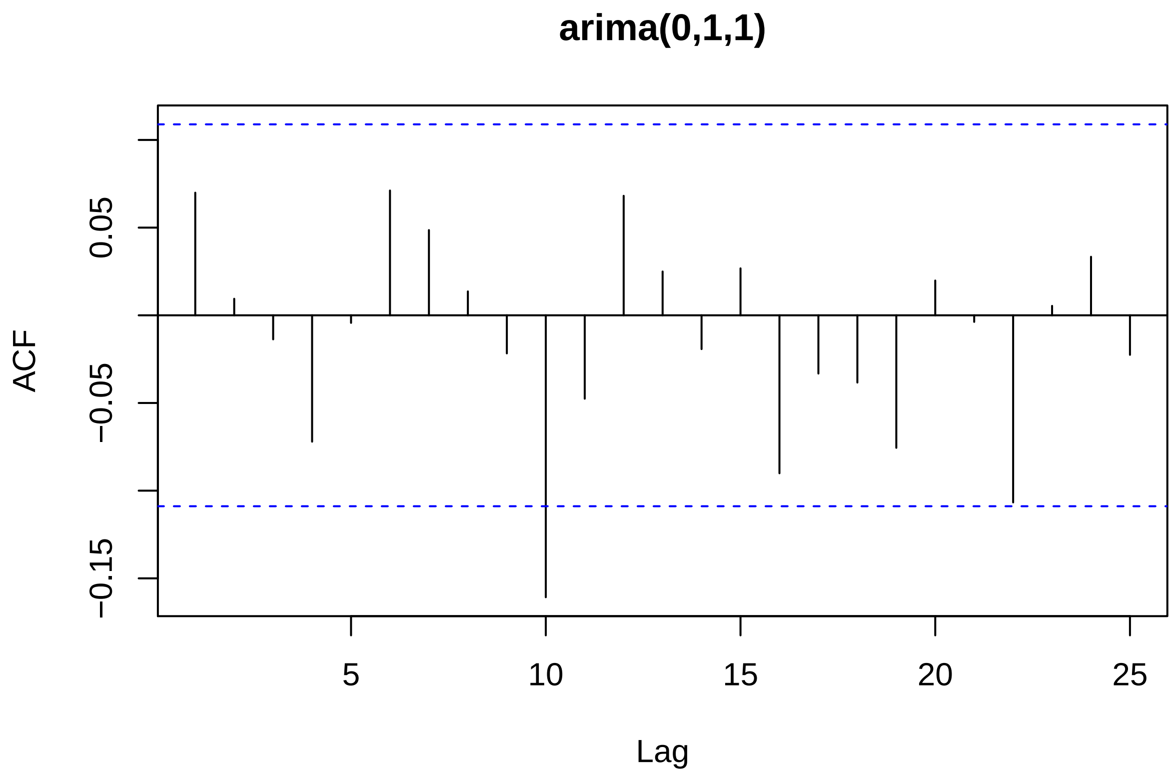




The arima(1,0,0) model has a lag problem with many lags being significant. The arima(0,1,1) model doesn't have any significant lags except for lag100 and the Ljung-Box test gives a good result of failing to reject the null hypothesis of not showing lack of fit. We will only move forward with the arima(0,1,1) model from this point on.

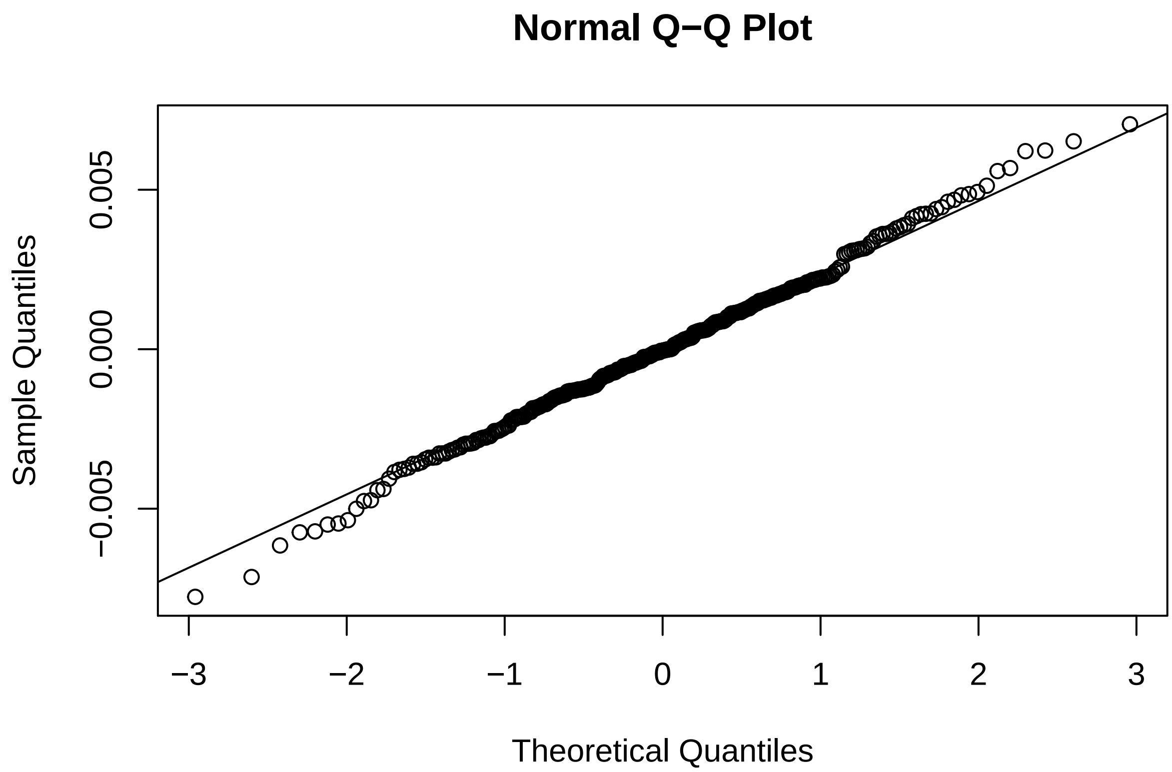



```
##
##  Box-Ljung test
##
## data:  residuals from  model2
## X-squared = 52.512, df = 11, p-value = 2.201e-07
```

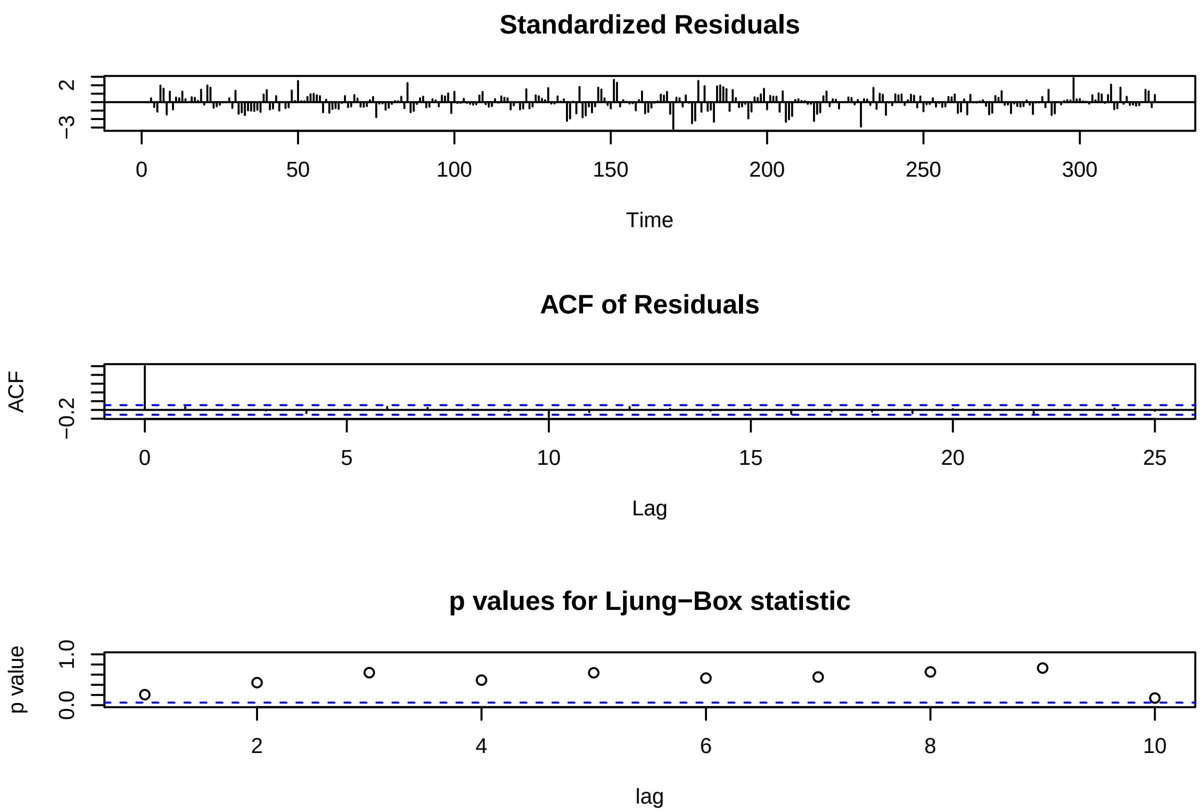



```
##
##  Box-Ljung test
##
## data:  residuals from  model3
## X-squared = 17.081, df = 11, p-value = 0.1055
```

From the following outputs, we can conclude that the arima(0,1,1) model fits well - normality is good, the residuals are also good, and the p-values are acceptable.




```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(model3)
## W = 0.99689, p-value = 0.791
```

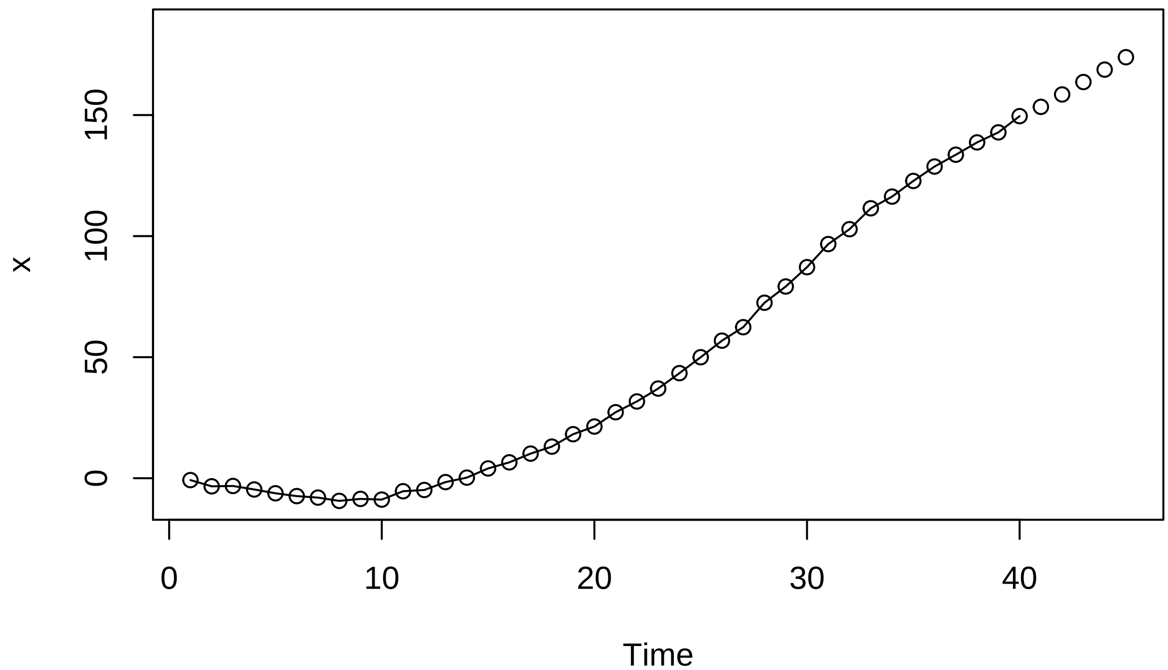


9.16

a)

```
##
## Call:
## arima(x = series2, order = c(0, 2, 2))
##
## Coefficients:
##           ma1     ma2
##       -0.9462  0.8548
## s.e.   0.1342  0.2646
##
## sigma^2 estimated as 1.299:  log likelihood = -60.36,  aic = 124.71
```

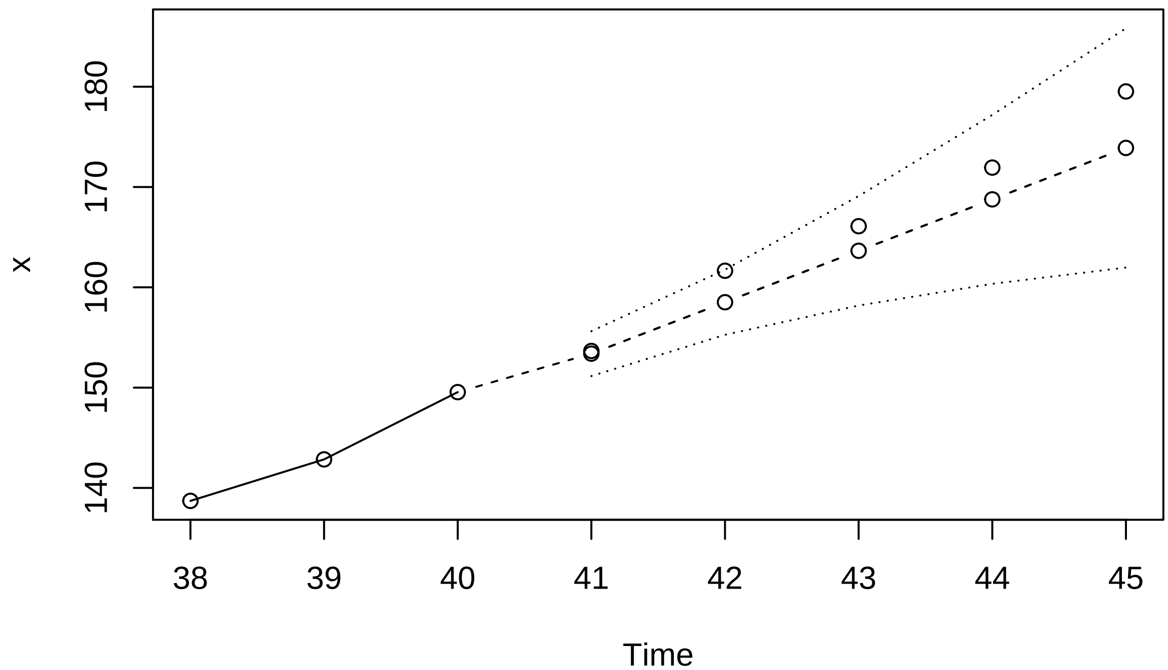
b) The forcasts are linear when the rest of the plot isn't.



c) The forecast appears to be underestimating the actual values.

```
## Time Series:
## Start = 41
## End = 45
## Frequency = 1
##      actual forecast
## 41 153.6557 153.3885
## 42 161.6516 158.5156
## 43 166.0946 163.6428
## 44 171.9380 168.7699
## 45 179.5304 173.8971
```

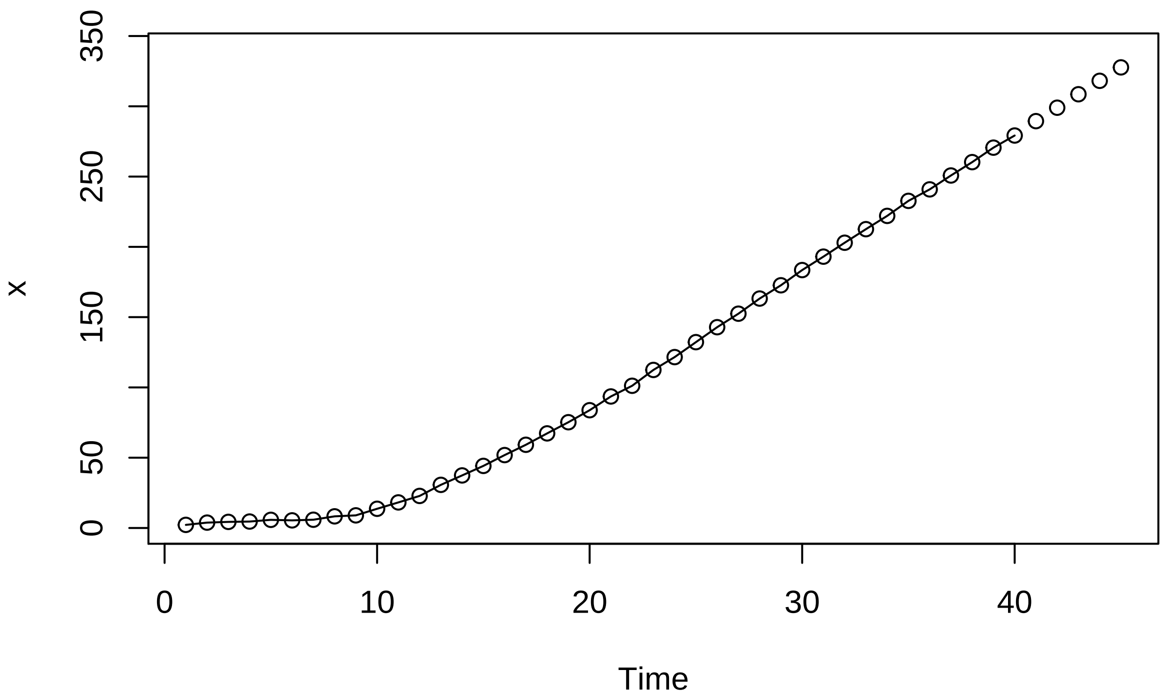
d) The actual values fall within the 95% CI except the point at 42 which appears to be on the CI boundary. From the actual values, time 42 is just below the upper CI bound.



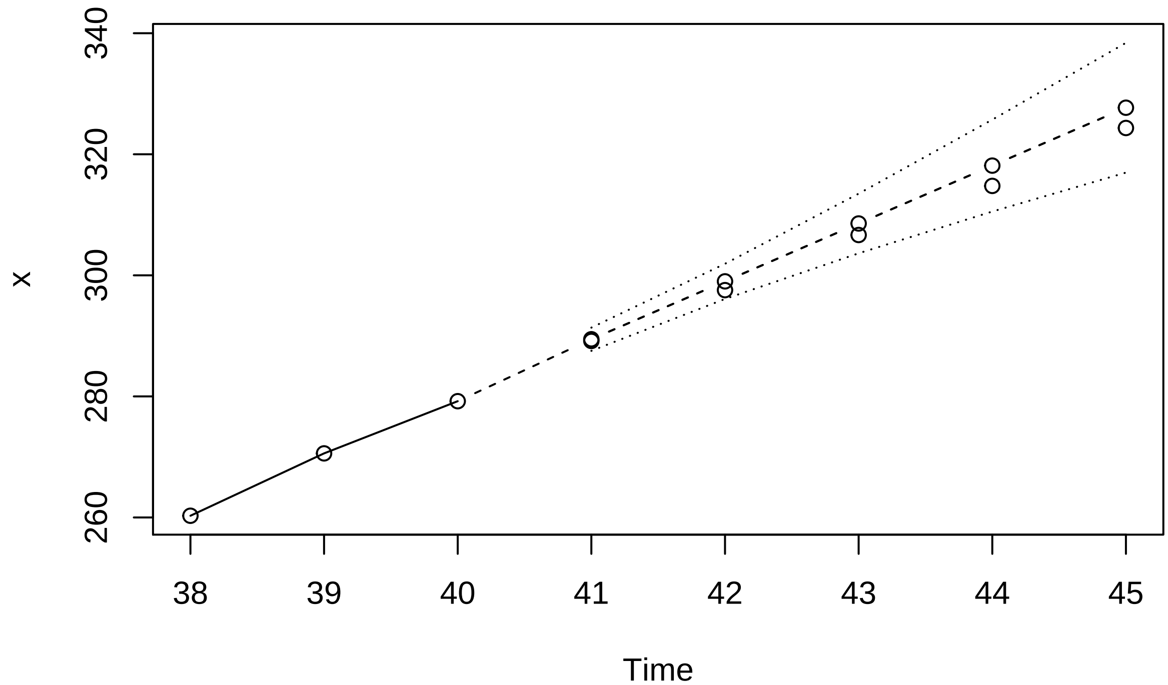

```
## Time Series:
## Start = 41
## End = 45
## Frequency = 1
##       lower   actual    upper
## 41 151.1538 153.6557 155.6232
## 42 155.2688 161.6516 161.7625
## 43 158.1863 166.0946 169.0992
## 44 160.3482 171.9380 177.1917
## 45 161.9705 179.5304 185.8236
```

e) At this different seed, the forecast is overestimating the actual values.

```
##
## Call:
## arima(x = series3, order = c(0, 2, 2))
##
## Coefficients:
##           ma1     ma2
##       -0.8626  0.7992
## s.e.   0.1249  0.1028
##
## sigma^2 estimated as 0.9572:  log likelihood = -54.24,  aic = 112.47
```



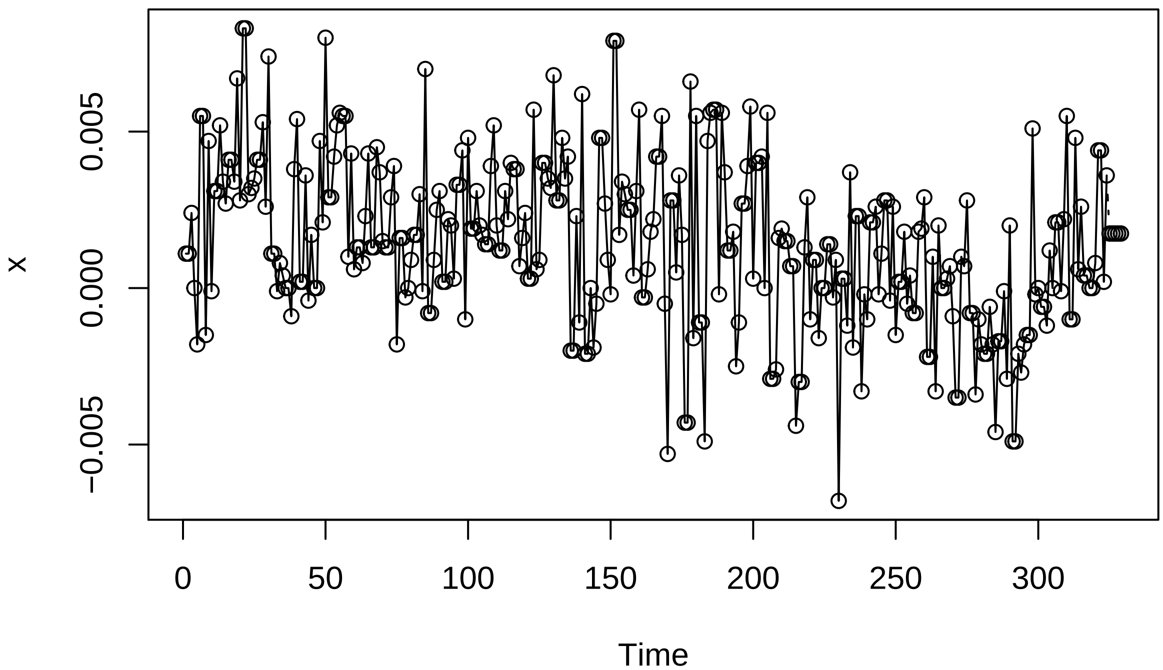

```
## Time Series:
## Start = 41
## End = 45
## Frequency = 1
##     actual2 forecast
## 41 289.1566 153.3885
## 42 297.5557 158.5156
## 43 306.6612 163.6428
## 44 314.7725 168.7699
## 45 324.3426 173.8971
```




```
## Time Series:
## Start = 41
## End = 45
## Frequency = 1
##      lower2  actual2   upper2
## 41 287.5230 289.1566 291.3583
## 42 296.0985 297.5557 301.9069
## 43 303.6402 306.6612 313.4893
## 44 310.5389 314.7725 325.7148
## 45 316.9714 324.3426 338.4064
```

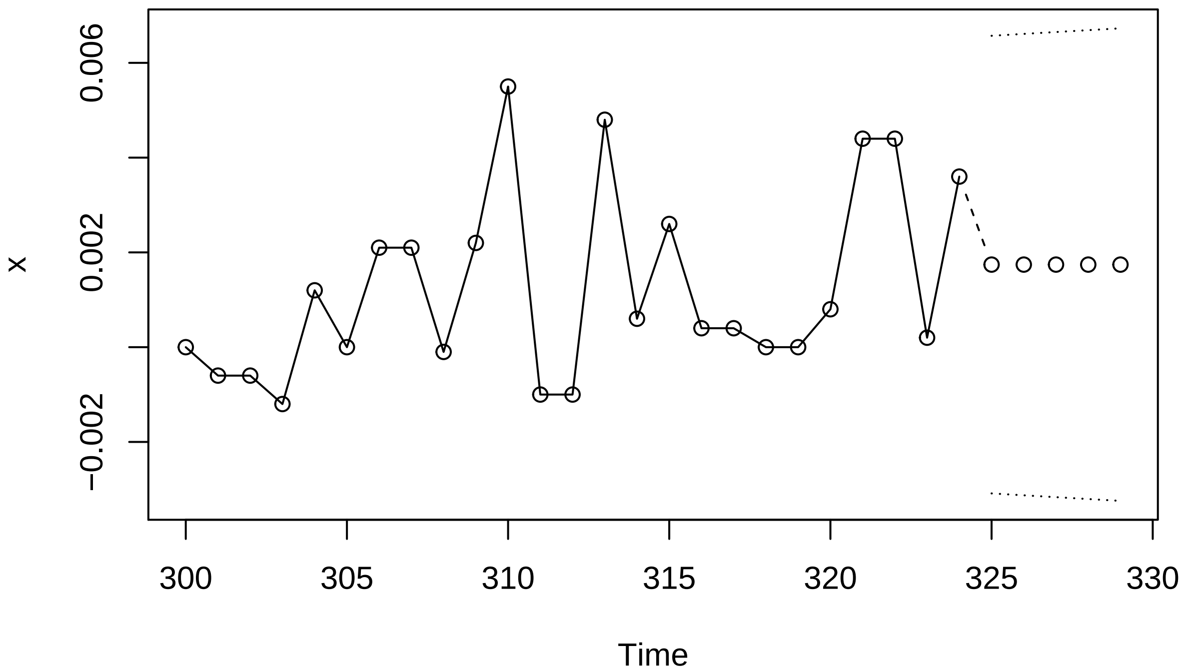
9.23

a)

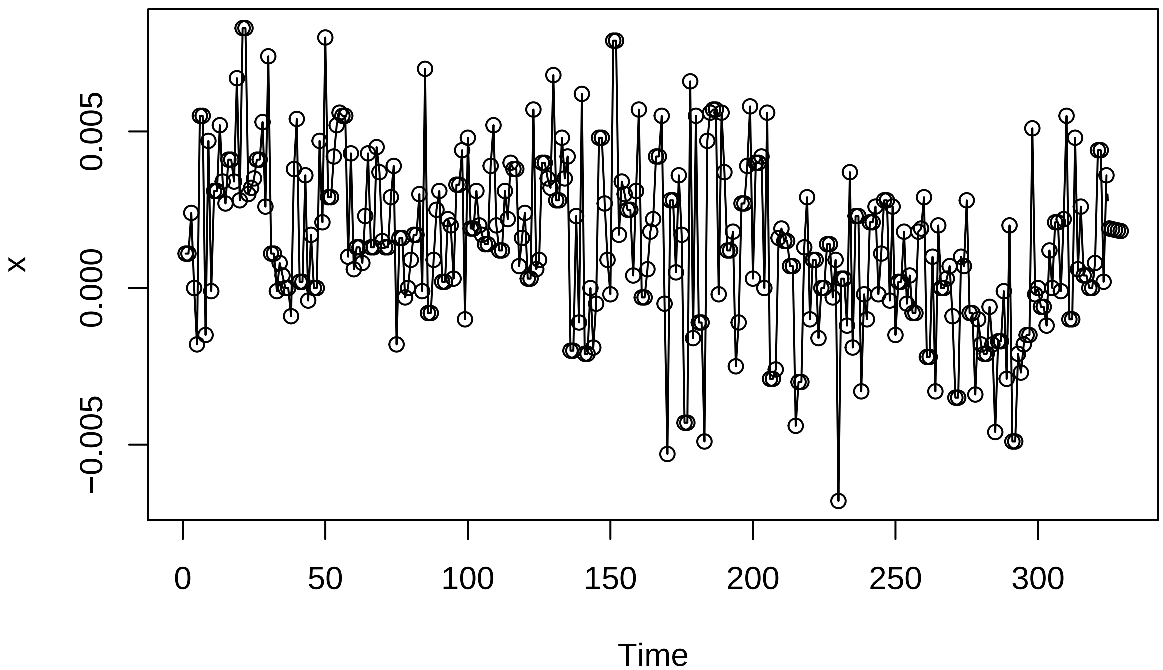



```
## Time Series:
## Start = 325
## End = 329
## Frequency = 1
##             lower   predicted       upper
## 325 -0.003086000 0.001742672 0.006571344
## 326 -0.003125839 0.001742672 0.006611183
## 327 -0.003165355 0.001742672 0.006650699
## 328 -0.003204555 0.001742672 0.006689898
## 329 -0.003243446 0.001742672 0.006728790
```

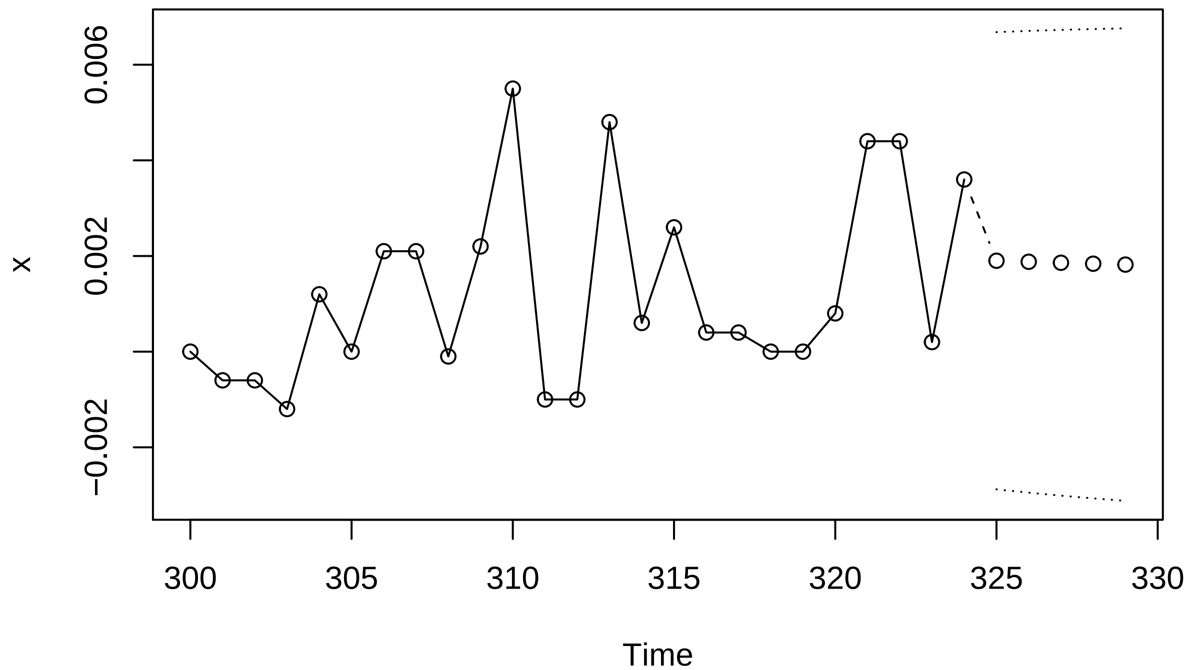
b) The forcasts are basically constant and the CI is very wide, likely due to the variance seen in previous values.



c) Both of these models produce very similar forecasts.




```
## Time Series:
## Start = 325
## End = 329
## Frequency = 1
##             lower   predicted       upper
## 325 -0.002878776 0.001901348 0.006681473
## 326 -0.002947994 0.001879444 0.006706881
## 327 -0.003010803 0.001858695 0.006728193
## 328 -0.003067889 0.001839041 0.006745972
## 329 -0.003119851 0.001820424 0.006760700
```


0.006
0.002
−0.002
x
300
305
310
315
320
325
330
Time

9.1 a) $\hat{Y}_t(1) = \mu + \phi(Y_t - \mu) = 10.8 + (-0.5)(12.2 - 10.8) = 10.1$

b) $\hat{Y}_t(2) = \mu + \phi(\hat{Y}_t(1) - \mu) = 10.8 + (-0.5)(10.1 - 10.8) = 11.15$

$\hat{Y}_t(2) = \mu + \phi^2(Y_t - \mu) = 10.8 + (-0.5)^2(12.2 - 10.8) = 11.15$

c) $\hat{Y}_t(10) = \mu + \phi^{10}(Y_t - \mu) = 10.8 + (-0.5)^{10}(12.2 - 10.8) = 10.801$

9.2 a) $Y_t = 5 + 1.1Y_{t-1} - 0.5Y_{t-2} + e_t \quad \sigma_e^2 = 2$

$\hat{Y}_t(\ell) = \phi_1 \hat{Y}_t(\ell-1) + \phi_2 \hat{Y}_t(\ell-2) + \ldots$

$\hat{Y}_{2008} = \hat{Y}_{2007}(1) = 5 + 1.1Y_{2007} - 0.5Y_{2006} = 5 + 1.1(10) - 0.5(11) = 10.5$ MILLION

$\hat{Y}_{2009} = \hat{Y}_{2007}(2) = 5 + 1.1\hat{Y}_{2008} - 0.5Y_{2007} = 5 + 1.1(10.5) - 0.5(10) = 11.55$ MILLION

b) $\left.\begin{array}{l} \psi_0 = 1 \\ \psi_1 - \phi_1\psi_0 = 0 \\ \psi_j - \phi_1\psi_{j-1} - \phi_2\psi_{j-2} \cdots = 0 \end{array}\right\}$ for AR(2)

$\psi_1 - \phi_1\psi_0 = 0$

$\psi_1 - \phi_1(1) = 0$

$\psi_1 = \phi_1 = 1.1$

c) $\hat{Y}_t(\ell) \pm z_{1-\alpha/2}\sqrt{Var(e_t(\ell))}$

$Var(e_t(\ell)) = \sigma_e^2 \sum^{\ell-1} \psi_j^2$

$\alpha = 0.05$

$\hat{Y}_{2007}(1) \pm z_{0.95}\sqrt{\sigma_e^2}$

$var(e_t(1)) = \sigma_e^2$

$10.5 \pm 2.575\sqrt{2}$

10.5 ± 3.6416

$(6.8584, 14.1416)$ MILLION

d) $Y_{2008} = 12$

$\hat{Y}_{2009} = \hat{Y}_{2008}(1) = 5 + 1.1Y_{2008} - 0.5Y_{2007} = 5 + 1.1(12) - 0.5(10) = 13.2$ MILLION

10.2 a) $(1 - 1.6x + 0.7x^2)(1 - 0.8x^{12})$

$\phi_1 = 1.6 \quad \phi_2 = -0.7 \quad \Theta_1 = 0.8$

$\phi_1 + \phi_2 = 1.6 - 0.7 = 0.9 < 1 \qquad \phi_2 - \phi_1 = -0.7 - 1.6 = -2.3 < 1 \qquad |\phi_2| = 0.7 < 1$

$\therefore$ STATIONARY

b) $ARIMA(2,0,0)\times(1,0,0)_{12}$

10.3 $Y_t = a + bt + S_t + X_t$ $\quad$ ARIMA$(p,0,q)\times(P,1,Q)_s$

$W_t = Y_t - Y_{t-s} = (a + bt + S_t + X_t) - (a + b(t-s) + S_{t-s} + X_{t-s})$

$= bs + S_t - S_{t-s} + X_t - X_{t-s} = bs + \nabla_s X_t$

ARIMA$(p,0,q)\times(P,0,Q)_s$

10.5 a) $Y_t = 0.5Y_{t-1} + Y_{t-4} - 0.5Y_{t-5} + e_t - 0.3e_{t-1}$

$Y_t - Y_{t-4} = 0.5Y_{t-1} - 0.5Y_{t-5} + e_t - 0.3e_{t-1}$

$= 0.5(Y_{t-1} - Y_{t-5}) + e_t - 0.3e_{t-1}$

$\Phi = 0.5$ $\quad \theta_1 = 0.3$ $\quad$ ARIMA$(1,0,1)\times(0,1,0)_4$

b) $Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0.5e_{t-1} - 0.5e_{t-12} + 0.25e_{t-13}$

$(Y_t - Y_{t-1}) - (Y_{t-12} - Y_{t-13}) = e_t - 0.5e_{t-1} - 0.5e_{t-12} + 0.25e_{t-13}$

$= e_t - 0.5(e_{t-1} + e_{t-12}) + (0.5)(0.5)e_{t-13}$

$\theta_1 = 0.5$ $\quad \Theta_1 = 0.5$ $\quad$ ARIMA$(0,1,1)\times(0,1,1)_{12}$