Homework 4

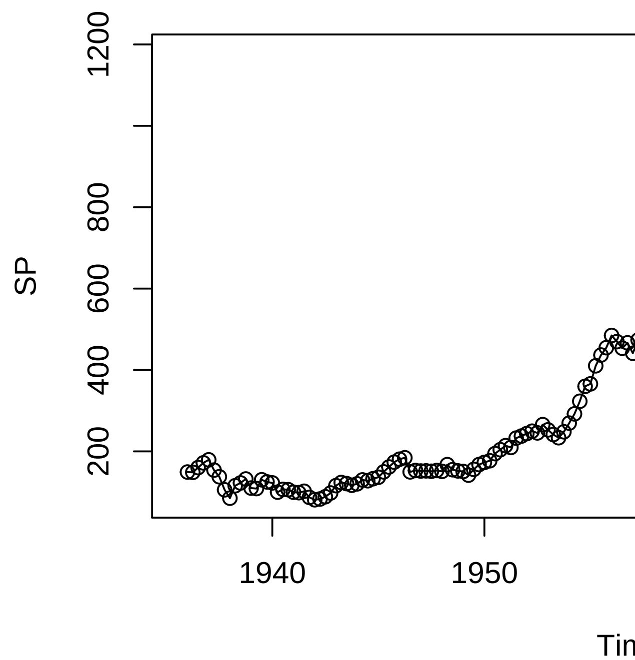
Deanna Springgay

4/2/2021

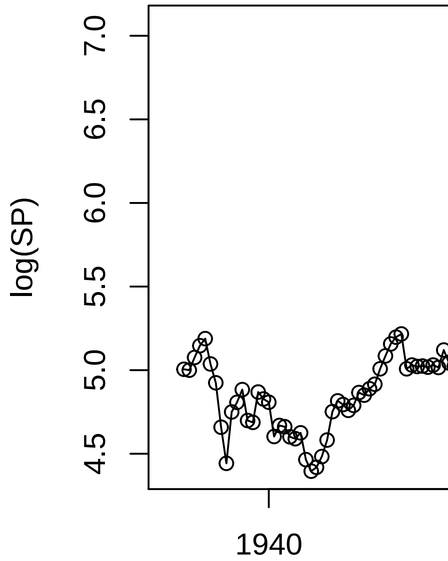
Problems

Written questions are attached at the end

5.12

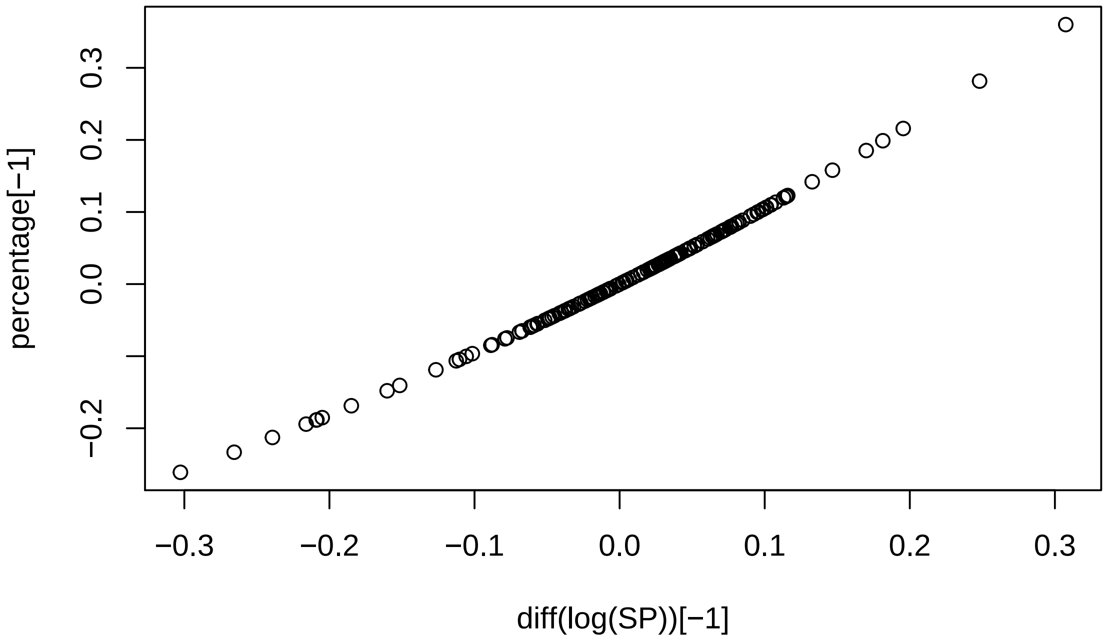


a) This time series shows a positive trend that levels off around 1970.



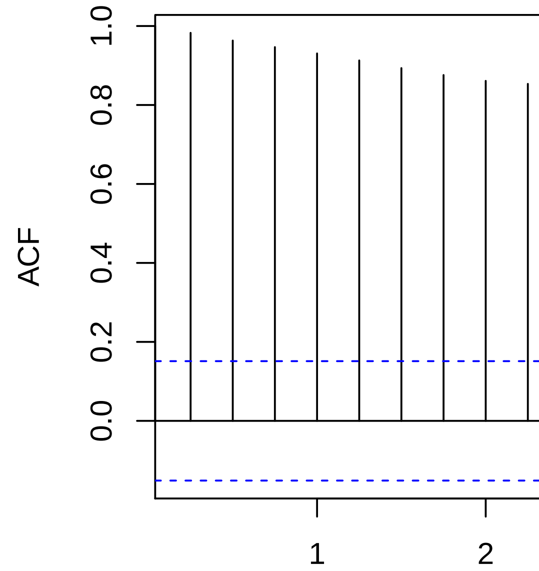
b) There isn't as much variation as the last plot, but the positive trend is still present.

c) Here the correlation is very strong at 99.6% between the fractional changes and the differences of loga-



rithms.

```
## [1] 0.9963475
```



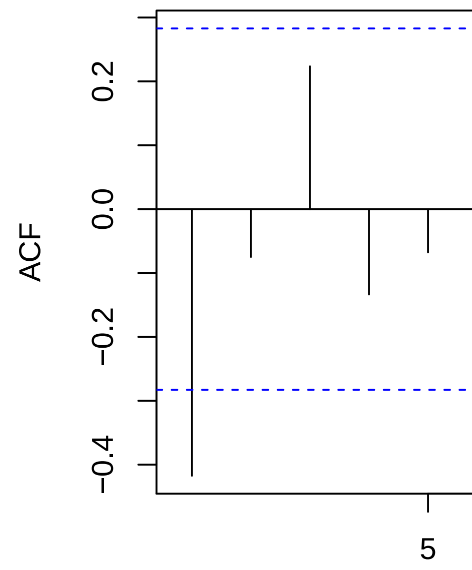
d) From the following information, we can conclude the data is nonstationary.

```
##
##  Augmented Dickey-Fuller Test
##
## data:  SP
## Dickey-Fuller = -3.2617, Lag order = 5, p-value = 0.0799
## alternative hypothesis: stationary
```

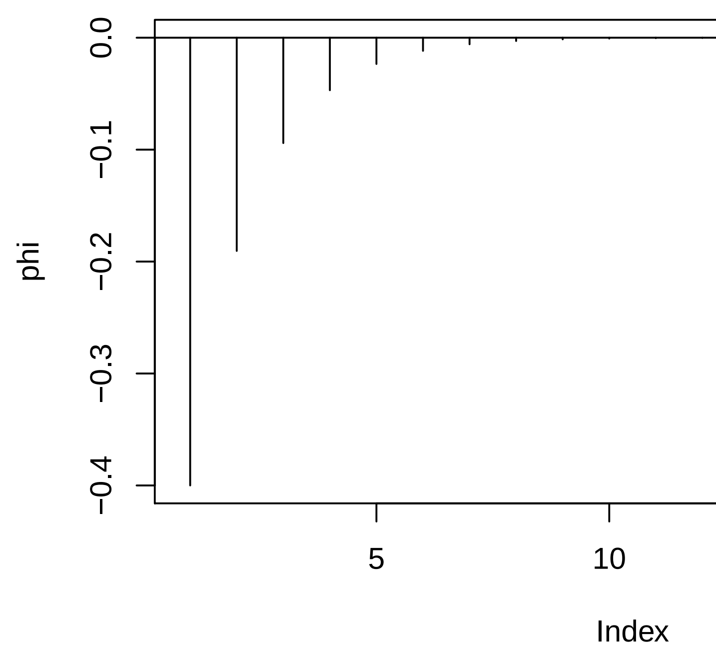
```
##
##  Phillips-Perron Unit Root Test
##
## data:  SP
## Dickey-Fuller Z(alpha) = -16.54, Truncation lag parameter = 4, p-value
## = 0.1677
## alternative hypothesis: stationary
```

6.26

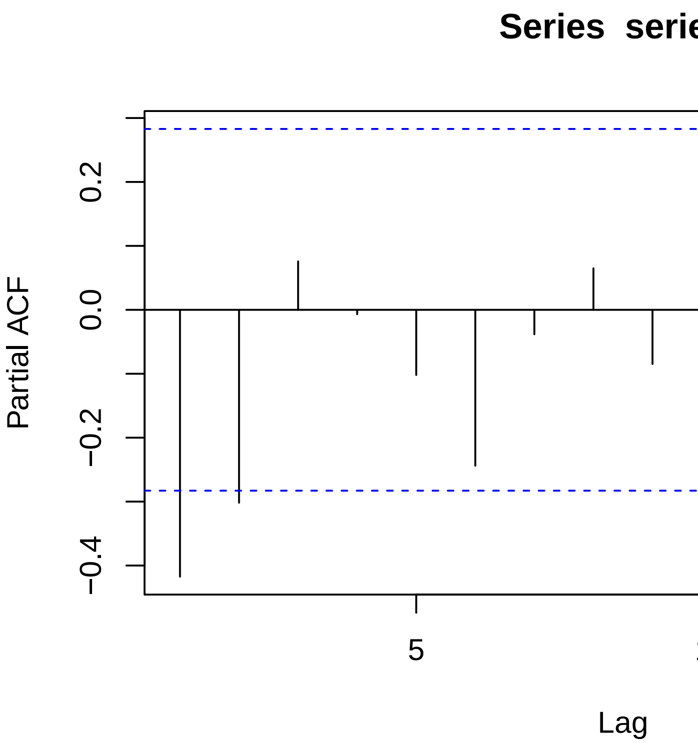
a) p1 = -(0.5)/(1+0.5^2) = -0.4 is the only nonzero autocorrelation.



b) ACF at lag 1 is fine, and there aren't any particularly strong correlations present.

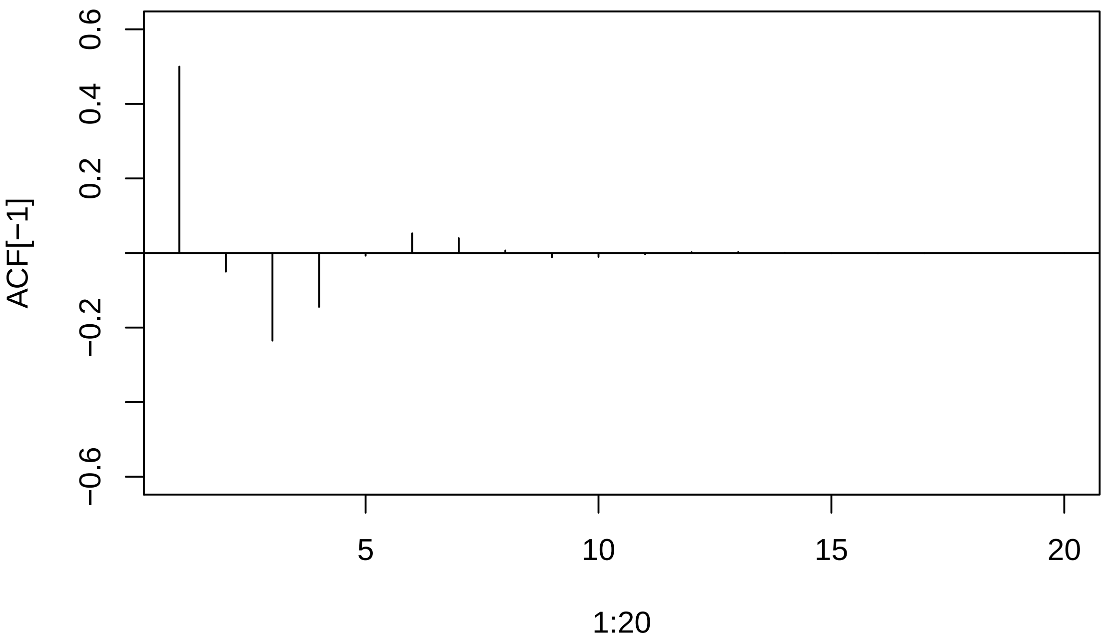


c) The correlations seem to disappear around lag 10 to lag 15.



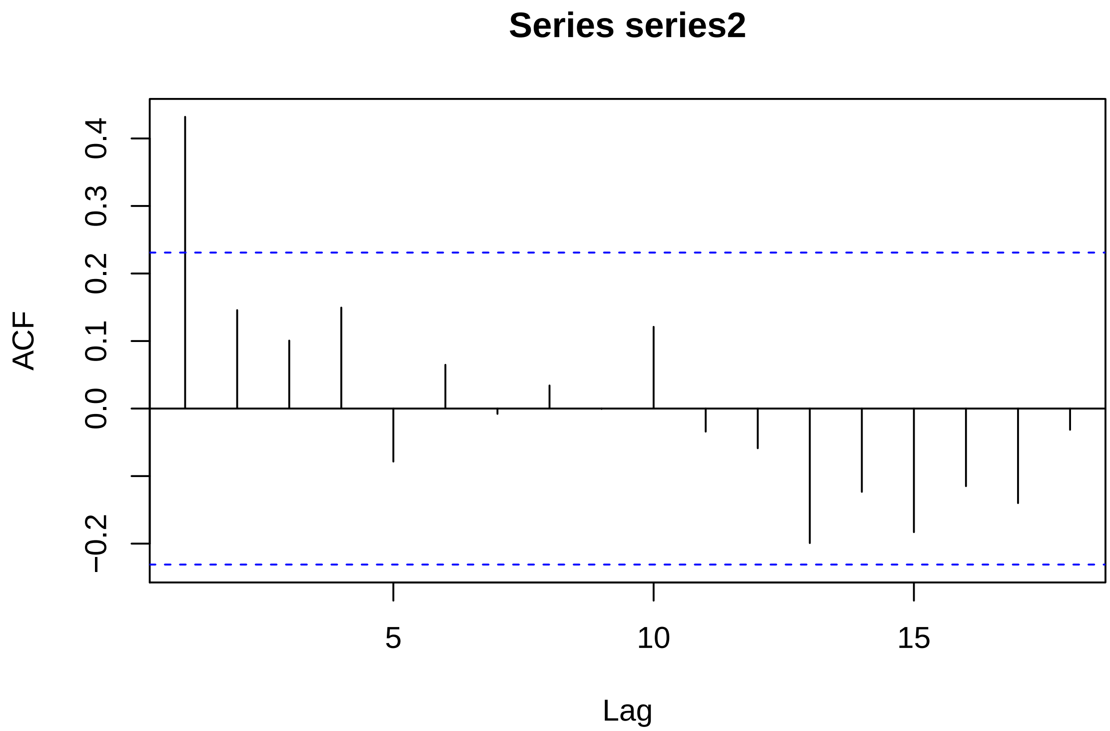
d) The first 2 lags match well, but the other lags are negligible.

6.27



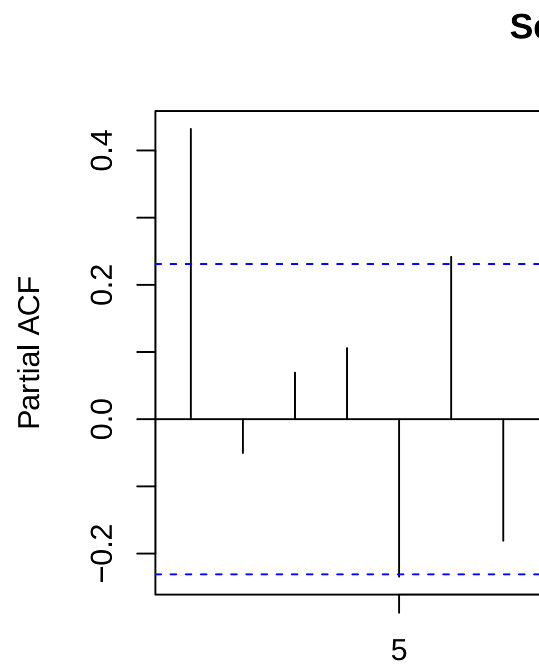
a)

b) Lag 1 matches well along with a wave pattern starting, but the pattern isn't dying out like expected from



the theoretical ACF.

c) phi11 = 0.5, phi22 = 0.7, and phi = 0 otherwise



d) The sample PACF doesn't match with the thoreticals as much as expected.

6.31

a) The Dickey-Fuller test leads us to reject nonstationarity.

```
## Warning in adf.test(series3, k = 0): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  series3
## Dickey-Fuller = -6.1909, Lag order = 0, p-value = 0.01
## alternative hypothesis: stationary
```

b) R chose k=4, and running this in the Dickey-Fuller test leads us to fail to reject nonstationarity.

```
##
## Call:
## ar(x = diff(series3))
##
## Coefficients:
##       1        2        3        4
## -0.6548  -0.6528  -0.4348  -0.2200
##
## Order selected 4  sigma^2 estimated as  0.8527
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  series3
## Dickey-Fuller = -2.3654, Lag order = 4, p-value = 0.427
```

```
## alternative hypothesis: stationary
```

c) At k=0, the DF test rejects nonstationarity. R chose r=6, and running the DF test again at k=6 we still reject nonstationarity though not to the same degree.

```
## Warning in adf.test(diff(series3), k = 0): p-value smaller than printed p-value
```

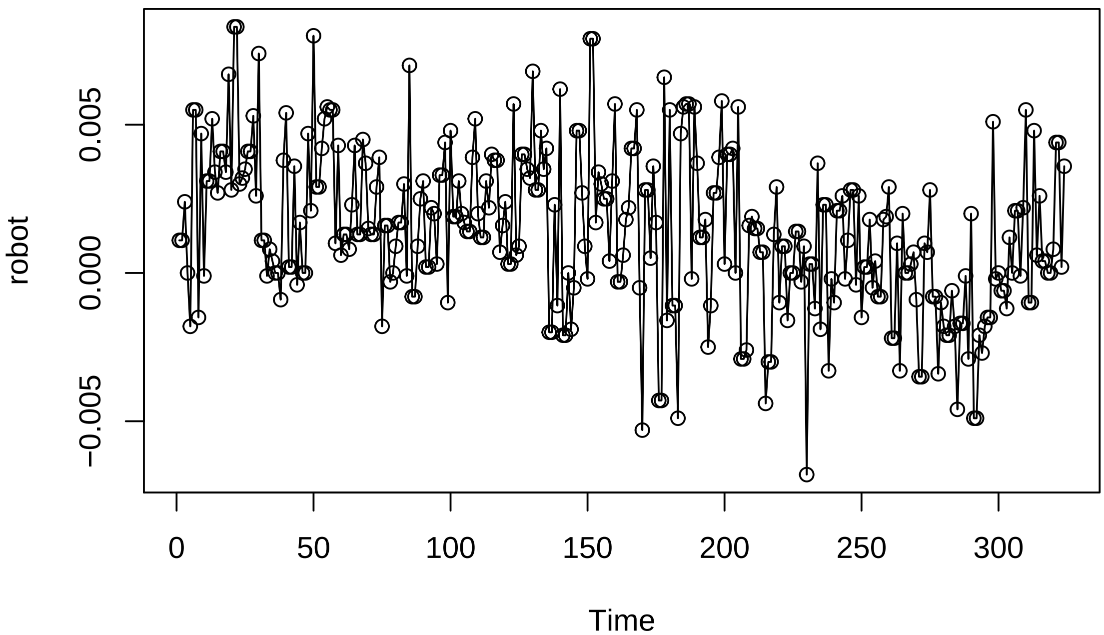
```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(series3)
## Dickey-Fuller = -10.62, Lag order = 0, p-value = 0.01
## alternative hypothesis: stationary
```

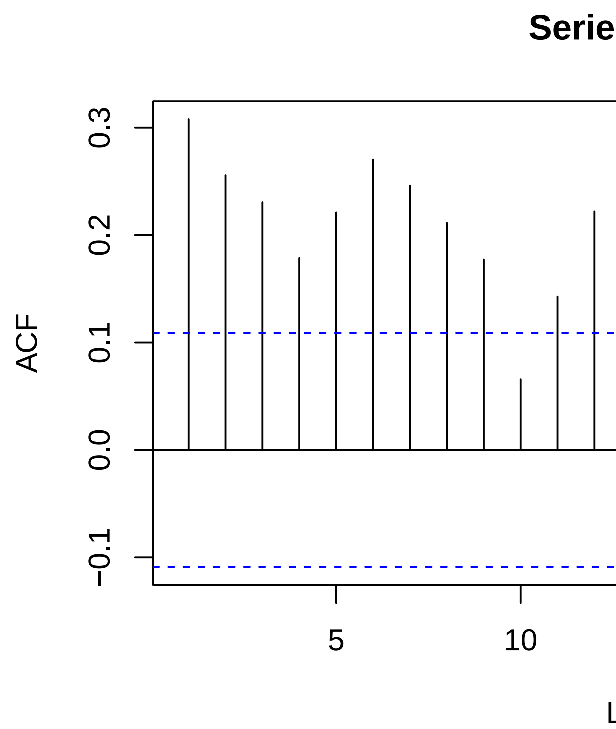
```
##
## Call:
## ar(x = diff(diff(series3)))
##
## Coefficients:
##       1        2        3        4        5        6
## -1.3819  -1.6359  -1.5413  -1.1884  -0.6975  -0.3245
##
## Order selected 6  sigma^2 estimated as  1.137
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(series3)
## Dickey-Fuller = -3.7062, Lag order = 6, p-value = 0.03185
## alternative hypothesis: stationary
```

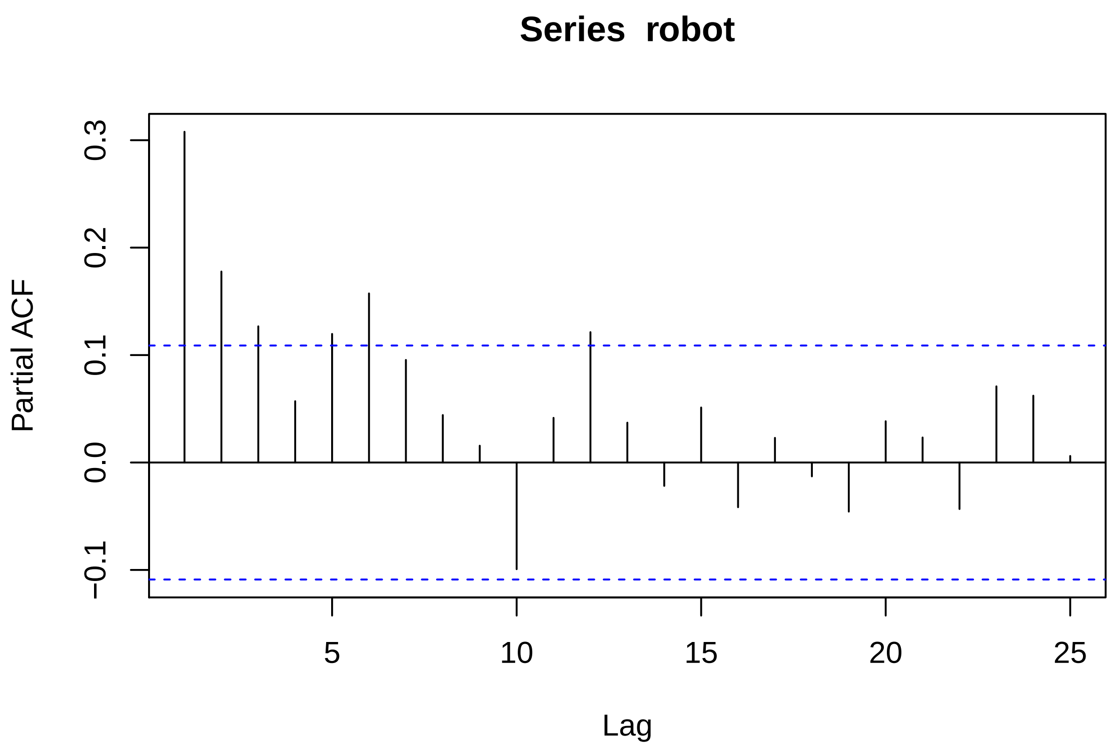
6.36

a) The data appears to be stationary, but we can also try a nonstationary model in case there is drift.





b) Neither of these plots are particularly helpful due to potential noise.



c) EACF indicates an ARIMA(1, 0, 1) model.

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 x x x x x x x x x o x  x  x  x
## 1 x o o o o o o o o o o  o  o  o
## 2 x x o o o o o o o o o  o  o  o
## 3 x x o o o o o o o o o  o  o  o
## 4 x x x x o o o o o o o  o  x  o
## 5 x x x o o o o o o o o  o  x  o
## 6 x o o o o x o o o o o  o  o  o
## 7 x o o x o x x o o o o  o  o  o
```

d) After checking a couple models, I would say this data would best be modeled by an ARIMA(1, 0, 1) model based off the output from the EACF since it had the lowest AIC, AICc, and BIC out of the 4 models I tested.

```
## Series: robot
## ARIMA(0,0,1) with non-zero mean
##
## Coefficients:
##          ma1    mean
##       0.2356  0.0015
## s.e.  0.0477  0.0002
##
## sigma^2 estimated as 6.699e-06:  log likelihood=1471.22
## AIC=-2936.45   AICc=-2936.37   BIC=-2925.1
```

```
## Series: robot
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##          ar1    mean
##       0.3076  0.0015
## s.e.  0.0528  0.0002
##
## sigma^2 estimated as 6.522e-06:  log likelihood=1475.54
## AIC=-2945.08   AICc=-2945   BIC=-2933.74
```

```
## Series: robot
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##          ar1      ma1    mean
##       0.9473  -0.8062  0.0015
## s.e.  0.0309   0.0609  0.0005
##
## sigma^2 estimated as 6.004e-06:  log likelihood=1489.3
## AIC=-2970.61   AICc=-2970.48   BIC=-2955.48
```

```
## Series: robot
## ARIMA(0,1,2)
##
## Coefficients:
##           ma1      ma2
##       -0.8088  -0.0930
## s.e.   0.0540   0.0594
##
```

```
## sigma^2 estimated as 6.057e-06:  log likelihood=1482.18
## AIC=-2958.36   AICc=-2958.29   BIC=-2947.03
```

5.1 a) $Y_t = 1 \cdot Y_{t-1} - 0.25 Y_{t-2} + e_t - 0.1 e_{t-1}$

$\phi_1 = 1 \quad \phi_2 = -0.25 \quad \theta_1 = 0.1 \longrightarrow$ ARIMA(2,0,1)

$\phi_1 + \phi_2 = 0.75 < 1 \quad \phi_2 - \phi_1 = -1.25 < 1 \quad |\phi_2| = 0.25 < 1$

$\therefore$ STATIONARY AND INVERTIBLE

b) $Y_t = 2Y_{t-1} - 1 \cdot Y_{t-2} + e_t$

$\phi_1 = 2 \quad \phi_2 = -1 \longrightarrow \phi_1 + \phi_2 = 1 \not< 1 \rightarrow$ NOT AR(2)

$Y_t - Y_{t-1} = (Y_{t-1} - Y_{t-2}) + e_t$

$\hookrightarrow$ AR(1), COEF $= 1$??

$Y_t - 2Y_{t-1} + Y_{t-1} = e_t \longrightarrow$ WN

$\therefore Y_t$ IS ARIMA(0,2,0)

c) $Y_t = 0.5 Y_{t-1} - 0.5 Y_{t-2} + e_t - 0.5 e_{t-1} + 0.25 e_{t-2}$

$\phi_1 = 0.5 \quad \phi_2 = -0.5 \quad \theta_1 = 0.5 \quad \theta_2 = -0.25$

$\phi_1 + \phi_2 = 0 < 1 \quad \phi_2 - \phi_1 = -1 < 1 \quad |\phi_2| = 0.5 < 1$

$\theta_1 + \theta_2 = 0.25 < 1 \quad \theta_2 - \theta_1 = -0.75 < 1 \quad |\theta_2| = 0.25 < 1$

$\therefore$ STATIONARY AND INVERTIBLE $\longrightarrow$ ARIMA(2,0,2)

6.12 $n = 100 \quad r_1 = -0.49 \quad r_2 = 0.31 \quad r_3 = -0.21 \quad r_4 = 0.11 \quad |r_k| < 0.004 \quad k > 4$

$2/\sqrt{100} = 0.2 \rightarrow$ MA(2) OR MA(3)? SIMPLER IS BETTER!

<u>TEST MA(2)</u>

$$\mathrm{Var}(r_k) = \frac{1}{n}\left[1 + 2\sum \rho_j^2\right] \qquad \mathrm{Var}(r_3) = \frac{1}{100}\left[1 + 3\left((-0.49)^2 + (0.31)^2\right)\right]$$

$$= 0.0167$$

$$r_3 / \mathrm{Var}(r_3) = -0.21 / 0.0167 = -1.62 \rightarrow \text{KEEP MA(2)!}$$

NO NEED TO TEST MA(3), MA(2) IS SIMPLER

6.13 $n = 121 \quad \hat{\phi}_{11} = 0.11 \quad \hat{\phi}_{22} = -0.6 \quad \hat{\phi}_{33} = 0.08 \quad \hat{\phi}_{44} = 0.08$

$2/\sqrt{121} = 0.18 \rightarrow$ AR(2) WORTH TESTING

6.19 a) LAG 1 OF SERIES A WILL BE STRONGLY POSITIVE SINCE 2 POINTS NEXT TO EACH OTHER DON'T HAVE A LARGE DIFFERENCE BETWEEN THEM. LAG 1 OF SERIES B WILL BE STRONGLY NEGATIVE SINCE EACH SUBSEQUENT POINT JUMPS TO THE OTHER SIDE OF THE MEAN TO A LARGE DEGREE.

b) LAG 2 OF SERIES A WILL ALSO BE POSITIVE SINCE 2 POINTS THAT ARE A JUMP AWAY FROM EACH OTHER ARE ALSO ON THE SAME SIDE OF THE MEAN. LAG 2 OF SERIES B IS STRONGLY POSITIVE SINCE THE 2 POINTS ARE ON THE SAME SIDE OF THE MEAN AND PRETTY CLOSE TO EACH OTHER.