lab-3-stat3925

Mason

14th March 2022

Question 1

We find the absolute value of the roots of $x^3 + 0.3x - 0.7x - 2 = 0$

```
roots = abs(polyroot(c(-2, -0.7, 0.3, 1)))
```

Since the roots are 1.3387633, 1.2222586 and 1.2222586 we see that all the roots are outside the unit circle

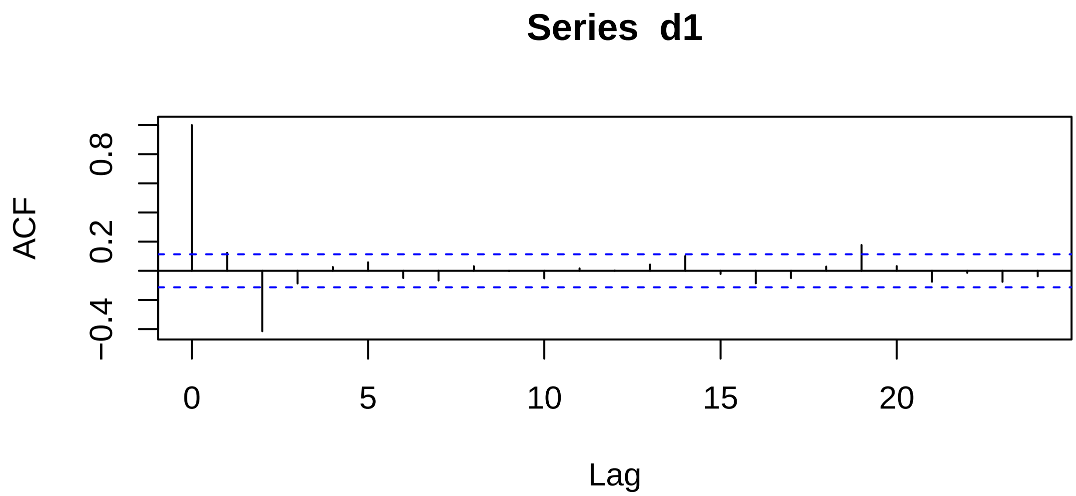
Question 2

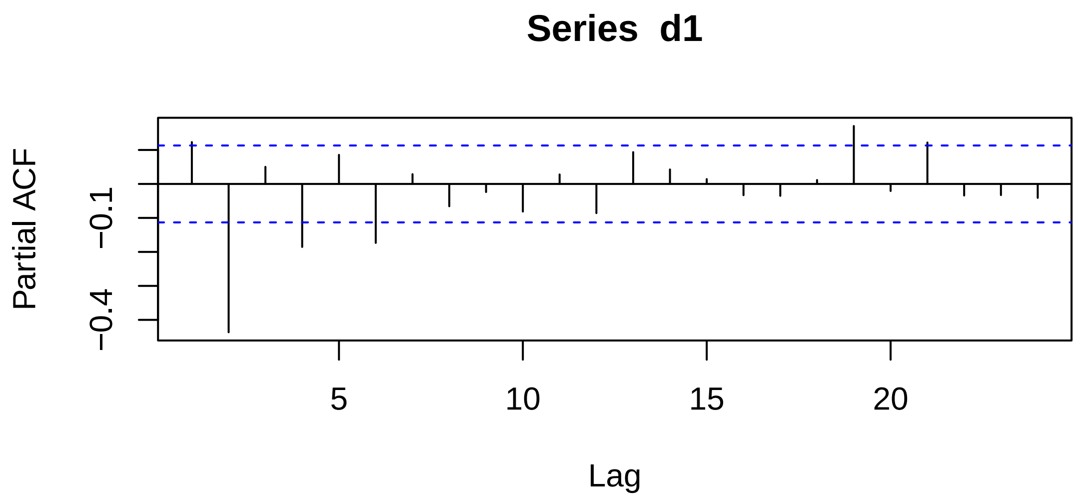
(i) We first begin by simulating 500 values from the model $X_t = Z_t + 0.6Z_{t-1} - 0.7Z_{t-2}$ where $Z_t \sim N(0, 1)$

```
d1 = arima.sim(list(ma = c(0.6, -0.7)), n = 500, sd = sqrt(1))
d1 = d1[201:500]
samp_mean = mean(d1)
samp_var = var(d1)
```

(ii) The sample values are close to the theoretical values.

```
par(mfrow = c(2, 1))
acf(d1)
pacf(d1)
```





(iii) There are significant spikes at lag-1 and lag-2 on average.

Question 3

Since the time series is modeled by: $X_t = Z_t + 0.6Z_{t-1} - 0.7Z_{t-2}$ where $Z_t \sim N(0, 1)$ we have that the $MA(2)$ polynomial is given by:

$$P(u) = 1 + 0.6u - 0.7u^2$$

```
quad_roots = abs(polyroot(c(1, 0.6, -0.7)))
quad_roots
```

```
## [1] 0.8411706 1.6983135
```

Since one root is inside and one root is outside the unit circle the $MA(2)$ model is not invertible.

Question 4

We simulate a series of length 450 values from

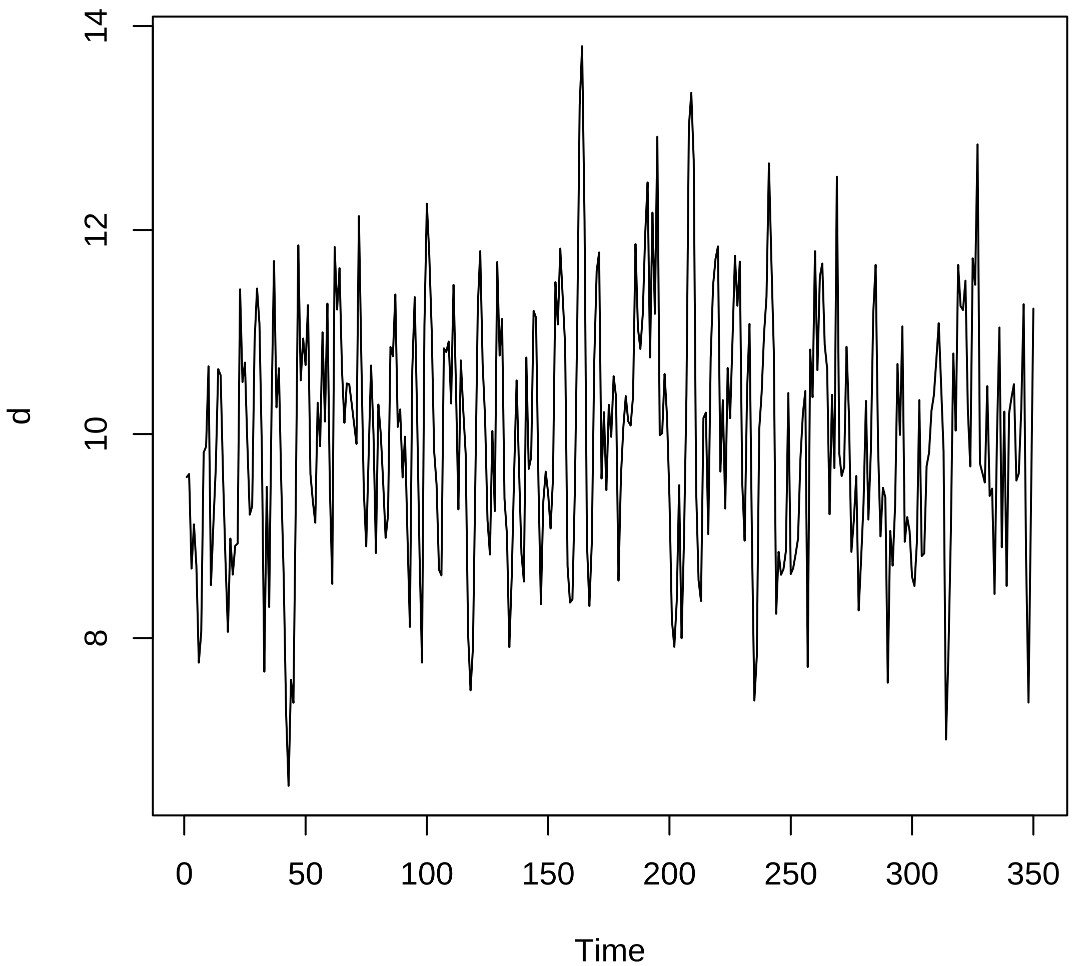
$$X_t = 10 + Z_t + 0.7Z_{t-1} + 0.5Z_{t-2}$$

Where $Z_t \sim NID(0,1)$

```
set.seed(123)
d = 10 + arima.sim(list(ma = c(0.7, 0.5)), n = 450, sd = sqrt(1))[101:450]
c(mean(d), var(d))
```

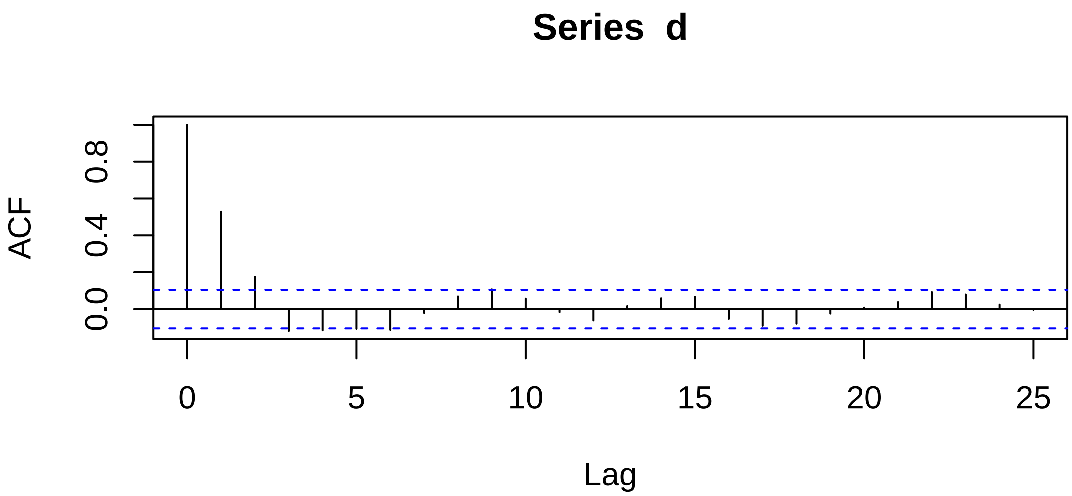
```
## [1] 9.977387 1.507868
```

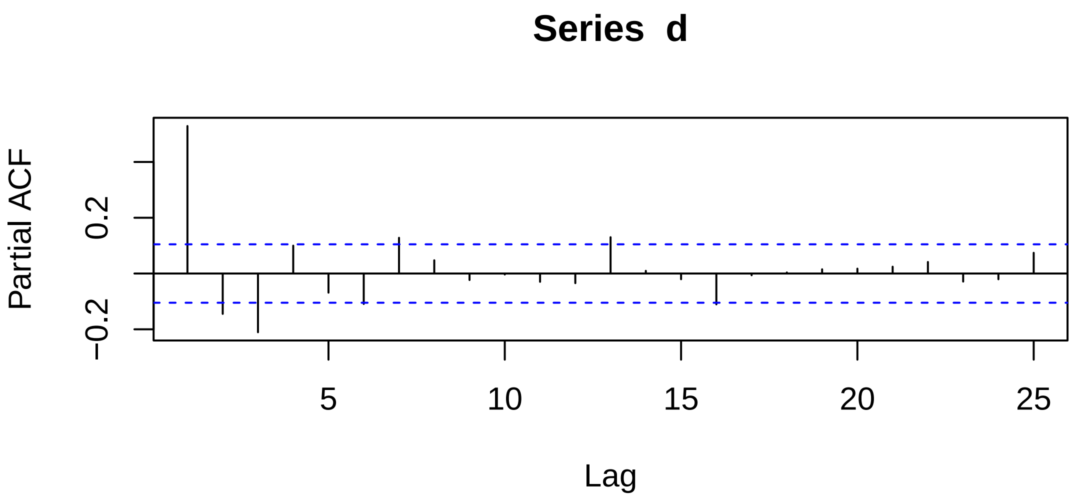
```
ts.plot(d)
```



The sample mean $\overline{x}$ is close to 10 because the theoretical mean $\mu = 10$

```
par(mfrow = c(2, 1))
r = acf(d)
p = pacf(d)
```





We have that the number of significant spikes in the ACF plot is 2. Hence, if we only had the ACF plot, we would be able to further investigate whether this was a $MA(2)$ model. We know that this is an $MA(2)$ model!

```
r[1:3]
```

```
##
## Autocorrelations of series 'd', by lag
##
##      1      2      3
##  0.529  0.175 -0.119
```

```
p[1:3]
```

```
##
## Partial autocorrelations of series 'd', by lag
##
##      1      2      3
##  0.529 -0.145 -0.211
```

Question 5

(a)

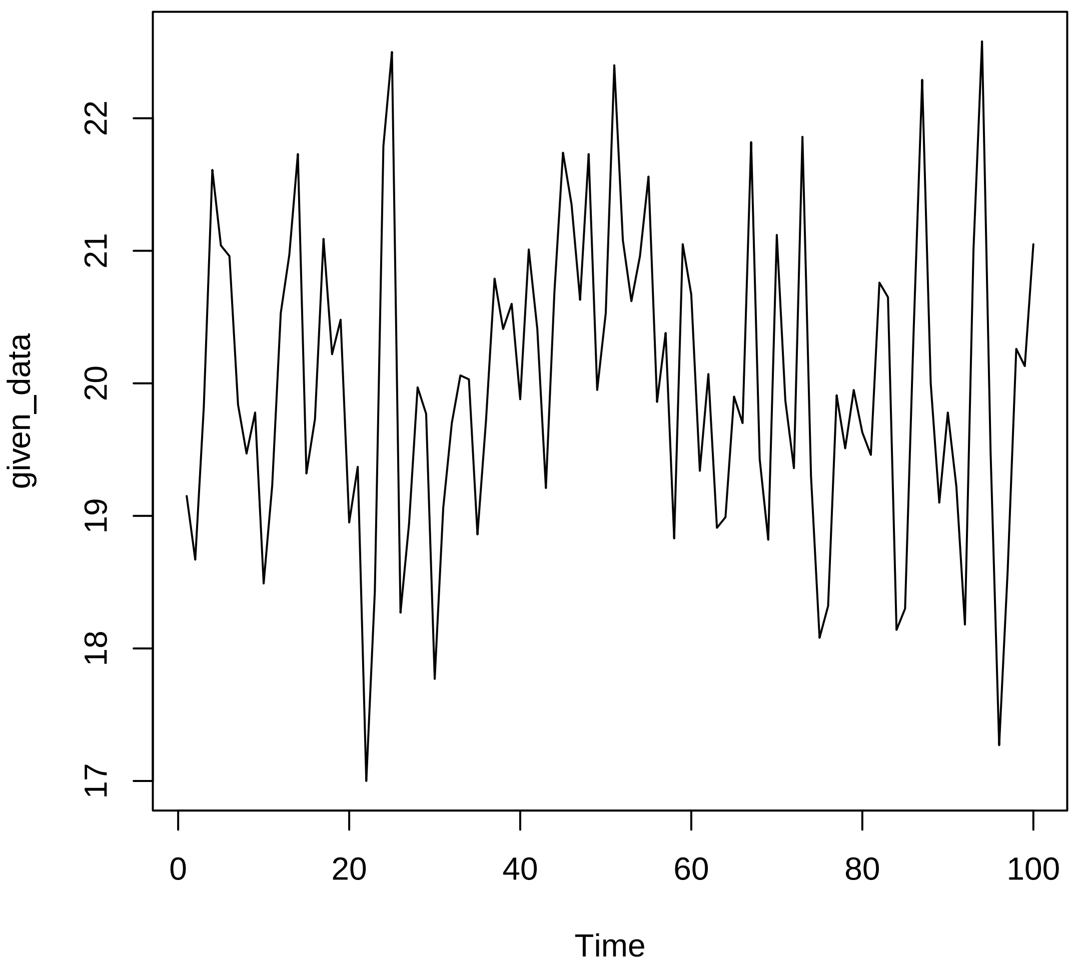
```
given_data = c(19.15, 18.67, 19.82, 21.61, 21.04, 20.96, 19.84, 19.47, 19.78, 18.49,
    19.23, 20.53, 20.97, 21.73, 19.32, 19.73, 21.09, 20.22, 20.48, 18.95, 19.37,
    17, 18.43, 21.79, 22.5, 18.27, 18.94, 19.97, 19.77, 17.77, 19.06, 19.7, 20.06,
    20.03, 18.86, 19.72, 20.79, 20.41, 20.6, 19.88, 21.01, 20.41, 19.21, 20.69, 21.74,
    21.35, 20.63, 21.73, 19.95, 20.53, 22.4, 21.08, 20.62, 20.96, 21.56, 19.86, 20.38,
    18.83, 21.05, 20.67, 19.34, 20.07, 18.91, 18.99, 19.9, 19.7, 21.82, 19.43, 18.82,
    21.12, 19.87, 19.36, 21.86, 19.3, 18.08, 18.32, 19.91, 19.51, 19.95, 19.63, 19.46,
    20.76, 20.65, 18.14, 18.3, 20.39, 22.29, 20, 19.1, 19.78, 19.22, 18.18, 21.01,
    22.58, 19.48, 17.27, 18.59, 20.26, 20.13, 21.05)

samp_mean = mean(given_data)
samp_var = var(given_data)
c(samp_mean, samp_var)
```

```
## [1] 19.99190  1.34243
```

(b)

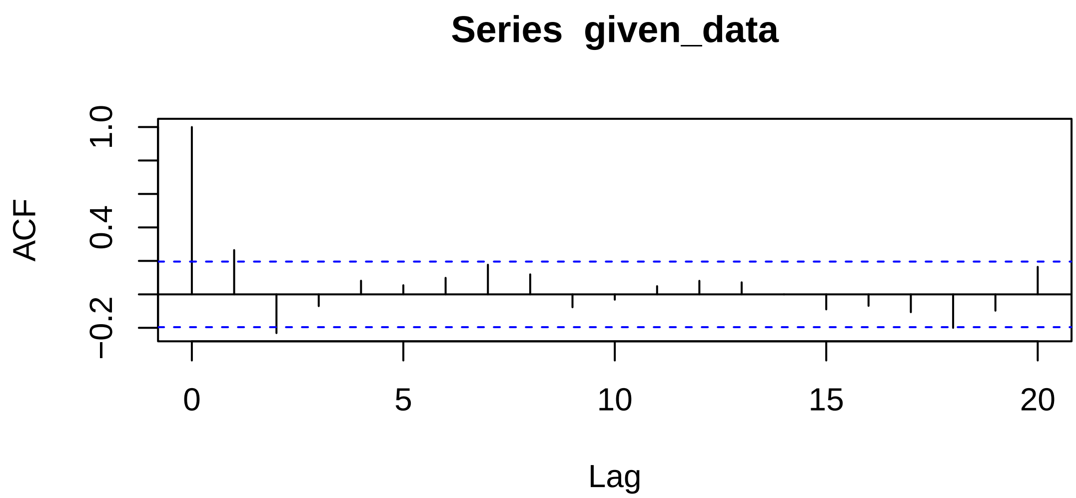
```
ts.plot(given_data)
```

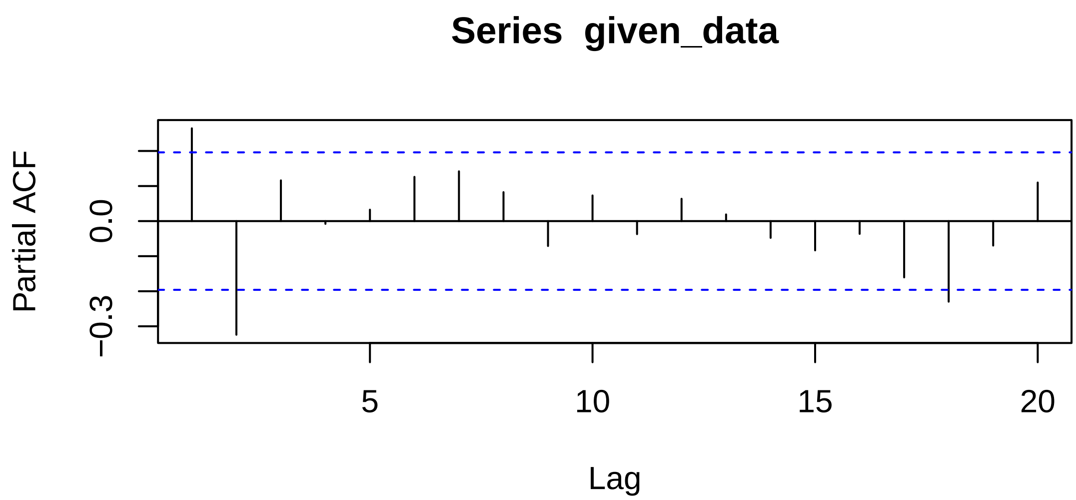


The time series plot looks to be stationary as the mean is roughly constant at a value of 20 and the variance also looks constant at about 1.5. This coincides with the values we obtained from before in (a).

(c)

```
par(mfrow = c(2, 1))
acf(given_data)
pacf(given_data)
```





The acf plot decays quickly so it reaffirms our idea that the time series is stationary.

(d) A model for this from the MA family is $MA(2)$ because we see 2 significant spikes at lag-1 and lag-2.