STAT 443 Lab 5

Derek Situ (62222807)

14/02/2022

Question 1

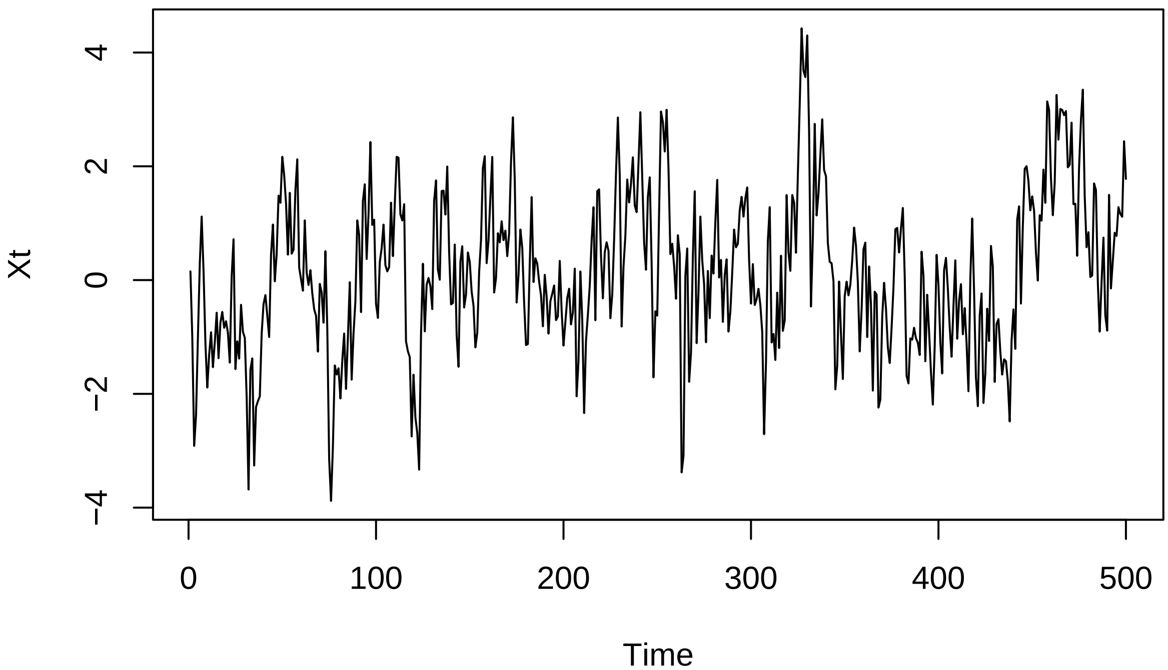
It is an AR(3) process.

Question 2

The acf of an AR(p) process should have slow decay and the pacf should cut off at lag p, so you can look for these to recognize it and determine its order p.

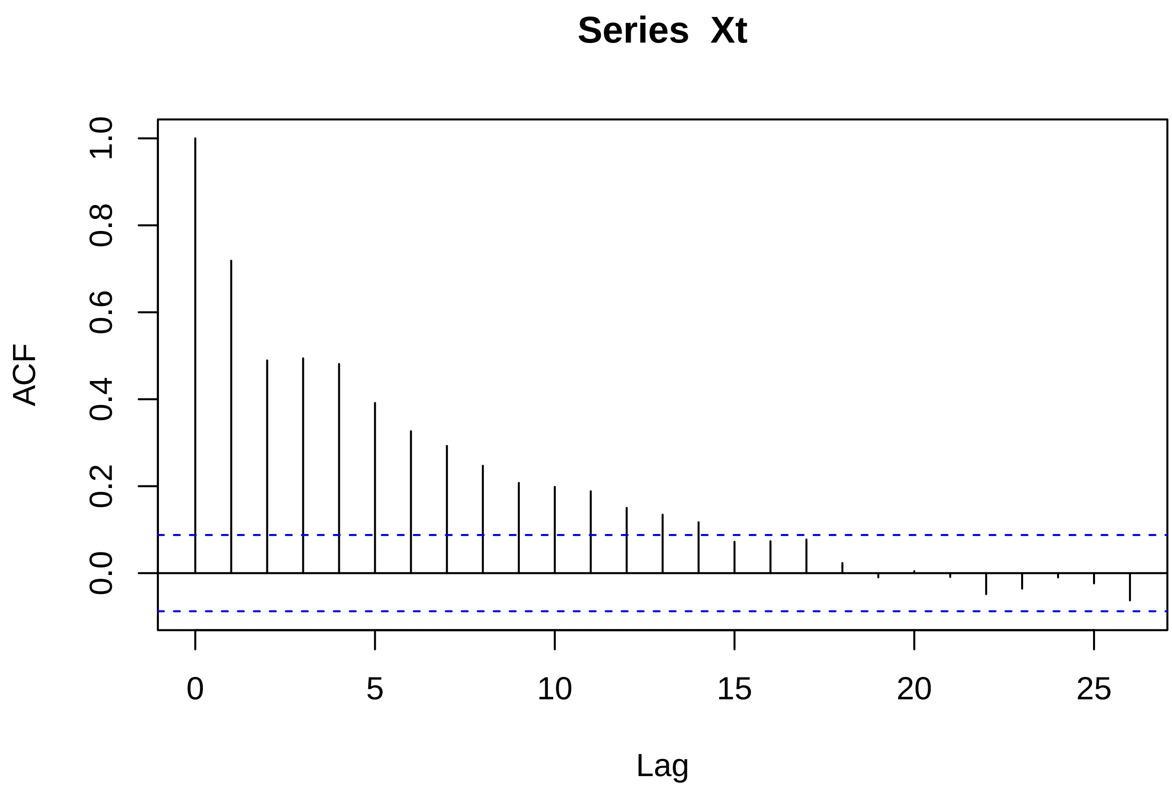
Question 3

```r
set.seed(123456)
Xt <- arima.sim(list(ar = c(0.8, -1/3, 0.6/sqrt(3))), n = 500, sd = sqrt(0.8))
plot(Xt)
```



Question 4

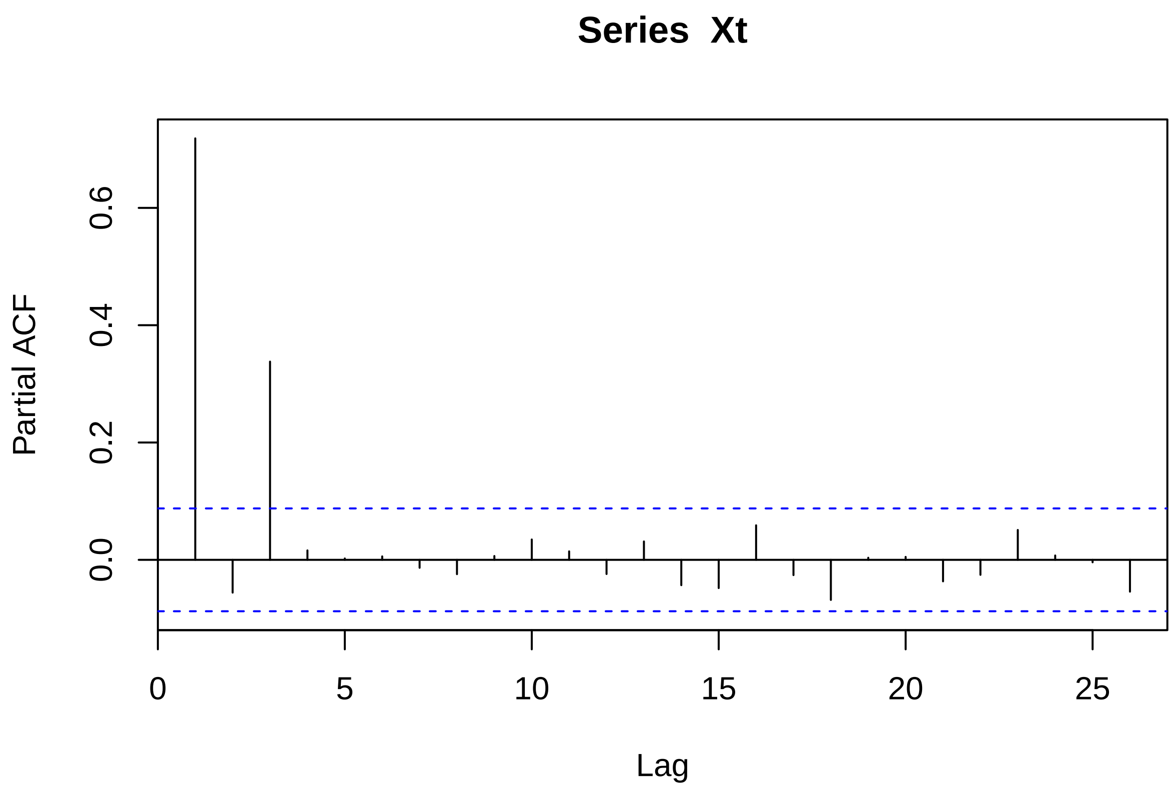
```
acf(Xt)
```



It behaves how I expect, as it decays slowly as is typical with AR processes.

Question 5

```
pacf(Xt)
```



For an AR(3) process we expect the pacf to cut off at lag 3, and we see that this is true for the sample pacf.

Question 6

Hypothesis test to see if the mean of the process $\mu = 0$:

$$H_0 : \mu = 0$$

Test statistic:

$$t = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Let's test at a 5% significance level, so we reject H_0 if $|t| > 1.96$.

```
t = mean(Xt)/(sd(Xt)/sqrt(length(Xt)))
t
```

```
## [1] 1.838177
```

Since $|t| < 1.96$, we cannot reject $H_0 : \mu = 0$, thus we will not include a mean to fit our model.

```
# fit an ARIMA(3, 0, 0) or AR(3) model to the simulated time series
arima_fit <- arima(Xt, order = c(3, 0, 0), include.mean = FALSE)

# Coefficients of the AR model
arima_fit$coef
```

```
##        ar1        ar2        ar3
##  0.7752982 -0.3103207  0.3423749
```

```
# Variance of Zt
arima_fit$sigma2
```

```
## [1] 0.7947373
```

Thus the estimated parameters are $\hat{\alpha}_1 = 0.775$, $\hat{\alpha}_2 = -0.310$, $\hat{\alpha}_3 = 0.342$, $\hat{\sigma}^2 = 0.795$.

Question 7

```
# fit an ARIMA(3, 0, 0) or AR(3) model to the simulated time series
arima_fit2 <- arima(Xt, order = c(3, 0, 0), method = "CSS",
                    include.mean = FALSE)

# Coefficients of the AR model
arima_fit2$coef
```

```
##        ar1        ar2        ar3
##  0.7728142 -0.3056298  0.3399697
```

```
# Variance of Zt
arima_fit2$sigma2
```

```
## [1] 0.7891485
```

The estimated parameters this time are $\hat{\alpha}_1 = 0.772$, $\hat{\alpha}_2 = -0.306$, $\hat{\alpha}_3 = 0.340$, $\hat{\sigma}^2 = 0.789$.

Let's compare the parameter estimates. The parameters used to simulate the series are in the column parameters, the estimated parameters from question (6) are in the column estimates1, and the estimated parameters from question (7) are in the column estimates2. The "diff" columns show the difference between the actual parameters and the parameter estimates. We can see that the model fitted with the default method produces closer estimates to the real parameters in this case.

```
parameters = c(0.8, -1/3, 0.6/sqrt(3), 0.8)
estimates1 = c(arima_fit$coef, arima_fit$sigma2)
estimates2 = c(arima_fit2$coef, arima_fit2$sigma2)
data.frame(parameters = parameters, estimates1 = estimates1,
           estimates2 = estimates2, diff1 = parameters - estimates1,
           diff2 = parameters - estimates2,
           row.names = c("alpha1", "alpha2", "alpha3", "sigma2"))
```

```
##        parameters estimates1 estimates2        diff1        diff2
## alpha1  0.8000000  0.7752982  0.7728142  0.024701780  0.027185829
## alpha2 -0.3333333 -0.3103207 -0.3056298 -0.023012660 -0.027703527
## alpha3  0.3464102  0.3423749  0.3399697  0.004035245  0.006440482
## sigma2  0.8000000  0.7947373  0.7891485  0.005262706  0.010851485
```