Celeste Tierre

PSTAT 174 HW #2

April 14, 2021

1. The dataset that exhibit statistically significant autocorrelation is (B) II only. I came up with that conclusion because it has the most amount of values in the outside of the portion compared to dataset I & III. whereas, dataset I has three valuer outside of the interval and mostly towards the middle line. While dataset only has one value outside the interval.

2(a) $X_t = Z_t - \frac{2}{3}Z_{t-1} - \frac{1}{3}Z_{t-2}$ $BX_t = X_{t-1}$ MA(2)

$= Z_t - \frac{2}{3}BZ_t - \frac{1}{3}B^2Z_{t-2}$

$= (1 - \frac{2}{3}B - \frac{1}{3}B^2)(Z_t)$

$1 - \frac{2}{3}B - \frac{1}{3}B^2 = 0$

$B^2 + 2B - 3 = 0$

$\rightarrow B^2 + 3B - B - 3 = 0$

$\rightarrow B^2 - B + 3B - 3 = 0$

$\rightarrow B(B-1) + 3(B-1) = 0$

$\rightarrow = (B-1)(B+3) = 0$

$B = 1$ $B = -3$

since the polynomial is known to be circle the moving process will be consistent. therefore, it will be stationary!
Hence, since $B=1$ is not know to be greater than 1 we can therefore state it's not invertible.

2(b) $X_t = \frac{2}{3}X_{t-1} + \frac{1}{3}X_{t-2} + Z_t$ AR(2)

$X_t - \frac{2}{3}X_{t-1} - \frac{1}{3}X_{t-2} = Z_t$

$X_t(1 - \frac{2}{3}B - \frac{1}{3}B^2) = Z_t$

$1 - \frac{2}{3}B - \frac{1}{3}B^2 = 0$

$\rightarrow B^2 + 2B - 3 = 0$

$\rightarrow B^2 - B + 3B - 3 = 0$

$B(B-1) + 3(B-1) = 0$

$(B-1)(B+3) = 0$

$B=1$ $B=-3$

AR(2) process isn't stationary because $B=1$ is not greater than 1 in magnitude. It is invertible because no conditions for AR to be invertible.

3(a) MA(3) $\theta_1 = 2$ $\theta_2 = .5$ $\theta_3 = -.1$

(i) mathematical equation of MA(3)

$$X_t = \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \theta_3 Z_{t-3} + Z_t$$

$$X_t = 2Z_{t-1} + .5Z_{t-2} - .1Z_{t-3} + Z_t$$

(ii) calculate autocorrelation func. at lags 1, 2, 3, 4, $\rho(1), \rho(2), \rho(3), \rho(4)$

lag=1 $\rho(1) = \dfrac{\theta_1 + \theta_1\theta_2 + \theta_2\theta_3}{1+\theta_1^2+\theta_2^2+\theta_3^2} \rightarrow \dfrac{(2)+(2)(.5)+(.5)(-.1)}{1+(2)^2+(.5)^2+(-.1)^2} = \dfrac{2.95}{5.26} = .561$

lag=2 $\dfrac{\sum_{i=0} \theta_i\theta_{i+1}}{\sum_{i=0}^{3} \theta_i^2} = \dfrac{\theta_2 + \theta_1\theta_3}{1+\theta_1^2+\theta_2^2+\theta_3^2} \rightarrow \dfrac{(.5)+(2)(-.1)}{1+(2)^2+(.5)^2+(-.1)^2} = \dfrac{.3}{5.26} = .057$

lag=3 $\dfrac{\sum_{i=0}^{0} \theta_i\theta_{i+1}}{\sum_{i=0}^{3} \theta_i^2} = \dfrac{\theta_3}{1+\theta_1^2+\theta_2^2+\theta_3^2} \rightarrow \dfrac{-.1}{1+(2)^2+(.5)^2+(-.1)^2} = \dfrac{-.1}{5.26} = -.019$

lag=4 = 0 b/c $\rho_x(k) = 0$ if $k > q$ $q = 3$

(3b) AR(1) $\phi_1 = -.5$

(i) write mathematical equation for AR(1)

$$X_t = \phi_1 X_{t-1} + Z_t$$

$$X_t = -.5X_{t-1} + Z_t$$

(ii) Calculate autocorrelation at 1, 2, 3, 4 $\rho(1), \rho(2), \rho(3), \rho(4)$

lag=1 $\rho(1) \rightarrow (-.5)^1 = -.5$

lag=2 $\rho(2)$ $(-.5)^2 = .25$

lag=3 $\rho(3) = (-.5)^3 = -.125$

lag=4 $\rho(4) = (-.5)^4 = .0625$

4 $X_t = 3 + Y + Z_t$

$Z_t \sim WN(0, \theta^2)$ $E[Z_t] = 0$ $E[Y] = 0$

$V[Z_t] = \theta^2$ $Var[Y] = \theta_Y^2$

Z & Y are independent.

$$E[X_t] = E[3 + Y + Z_t]$$
$$= 3 + E[Y] + E[Z_t]$$
$$= 3 + 0 + 0$$
$$E[X_t] = 3$$

$$Var[X_t] = Var[3 + Y + Z_t]$$
$$= 0 + Var[Y] + Var[Z_t] + 2cov(3, Y) + 2cov(Y, Z_t) + 2cov(3, Z_t)$$
$$= \theta_Y^2 + \theta^2$$

$var(X_t) < \infty$ $E|X_t|^2 < \infty$

$$cov(X_t, X_{t+h}) = cov(3 + Y + Z_t, 3 + Y + Z_{t+h})$$
$$= cov(Y, Y) + cov(Y, Z_{t+h}) + cov(Z_t, Y) + cov(Z_t, Z_{t+h})$$

$$cov(X_t, X_{t+h}) = \theta_Y^2 + 0 + 0 + 0$$

$$cov(X_t, X_{t+h}) = \begin{cases} \theta^2 + \theta_Y^2 & h = 0 \\ \theta_Y^2 & h \neq 0 \end{cases}$$

Autocovariance is $cov(X_t, X_{t+h})$ solution

Autocorrelation $= \dfrac{cov(X_t, X_{t+h})}{\sqrt{Var(X_t)Var(X_{t+h})}}$ $\dfrac{\theta_Y^2}{\theta_Y^2 + \theta_Z^2}$ $ACF = \begin{cases} 1 & h = 0 \\ \dfrac{\theta_Y^2}{\theta_Y^2 + \theta_Z^2} & h \neq 0 \end{cases}$

The process is stationary & autocov/autocor don't rely on it.

5 $X_t = Z_t + 2Z_{t-1} - 8Z_{t-2}$

(i) The model in this instance is known to be $ma(q)$ with $q=2$. Therefore, it is $ma(2)$

(ii) please refer to the pdf file o R of the following page

(iii) $$p(k) = \frac{\sum_{n=0}^{q-k} \theta_i \theta_{k+i}}{1+\theta_1^2 \cdots \theta_q^2}$$

$$p(2) = \sum_{i=0}^{2-n} \frac{\theta_i \theta_{2+i}}{1+\theta_1^2+\theta_2^2} = \frac{\theta_0 \theta_2 + 0}{1+\theta_1^2+\theta_2^2} = \frac{(1)(-8)}{1+(2)^2+(-8)^2} = \frac{-8}{69} = -.1159$$

PSTAT174_HW2

Celeste Herrera

4/14/2021

5. Let $X_t = Z_t + 2Z_{t-1} - 8Z_{t-2}$.

(i) Identify the model as the model as MA(q) or AR(p), specify q or p respectively.

Please refer to previous page

(ii) Is the model stationary and invertible? Explain fully and show calculations where needed. (Hint: review 4 from homework 1!)

```
polyroot(c(1,2,-8))
```

```
## [1] -0.25+0i  0.50+0i
```

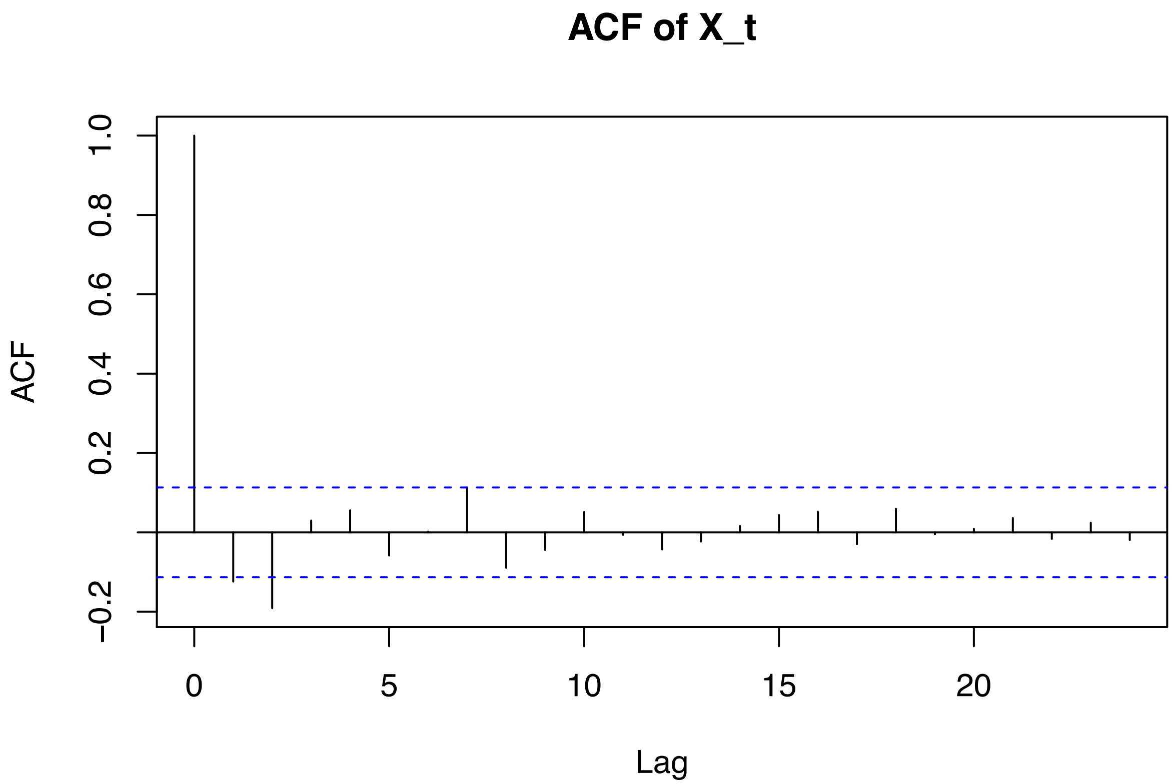
The values of our roots above which turned out to be -0.25+0i 0.50+0i. We see that both roots are within the unit circle not greater than 1. Therefore we would indicate that this MA(2) time series is not invertable. All moving average processes are stationtionary, so therefore time series above is stationary.

5 (iii) Find $\rho_X(2)$. Use R to simulate 300 values of $\{X_t\}$ and use your simulated values to plot sample acf. Compare your sample estimate of ρ_X (2) to its true value found by calculations. Redo this part using 10,000 simulated values of X_t.

```
#simulation of 300 values
value <- rnorm(300,0,1)

x_t<- filter(value, filter = c(1,2,-8), sides = 2, method = "convolution")

acf(x_t, main = "ACF of X_t", na.action = na.pass)
```




```
acf(x_t, lag.max = 2, plot = FALSE, na.action = na.pass)
```

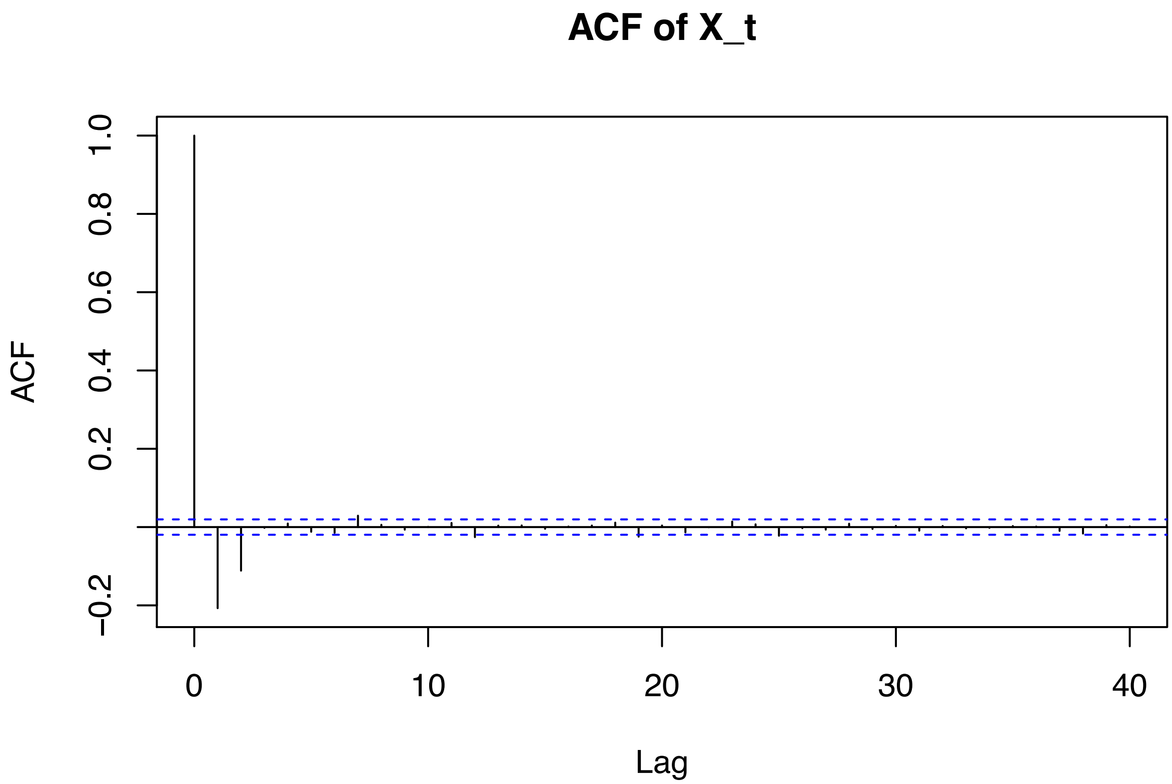
```
##
## Autocorrelations of series 'x_t', by lag
##
##      0      1      2
##  1.000 -0.124 -0.191
```

```
#simulation of 10,000 values

value_2 <- rnorm(10000,0,1)

x_t_2<- filter(value_2, filter = c(1,2,-8), sides = 2, method = "convolution")

acf(x_t_2, main = "ACF of X_t", na.action = na.pass)
```




```r
acf(x_t_2, lag.max = 2, plot = FALSE, na.action = na.pass)
```

```
##
## Autocorrelations of series 'x_t_2', by lag
##
##      0      1      2
##  1.000 -0.207 -0.111
```

(Also look at prior claulations on previous page)

For the above calculations of our ACF we find that our simulation for lag(2) when it is 300 values to be -.152 but when we had our values set to be at 10,000 whe had a lag(2) estimate at -0.110 which is to be a .04 difference between both values. Therefore we now have values close to lag(2).