Name: Jiacheng Niu
Course: Fintech 545

Problem 1
Remember from last week we discussed that skewness and kurtosis functions in statistical packages are often biased. Is your function biased? Prove or disprove your hypothesis.

Here is the result based on the code:

Sample skew mean: 0.00015319182838459739
Population skew: 0.0
t-statistic for skew: 0.6104981767907202
p-value for skew: 0.5416707083643297
Sample kurtosis mean: 3.000011427334777
Population kurtosis: 0.0
t-statistic for kurtosis: 6324.263372647042
p-value for kurtosis: 0.0
The p-value for skew is greater than 0.05, cannot reject the null hypothesis.
The p-value for kurtosis is less than 0.05, reject the null hypothesis.

So, the function of skewness is unbiased, and the function of kurtosis is biased.

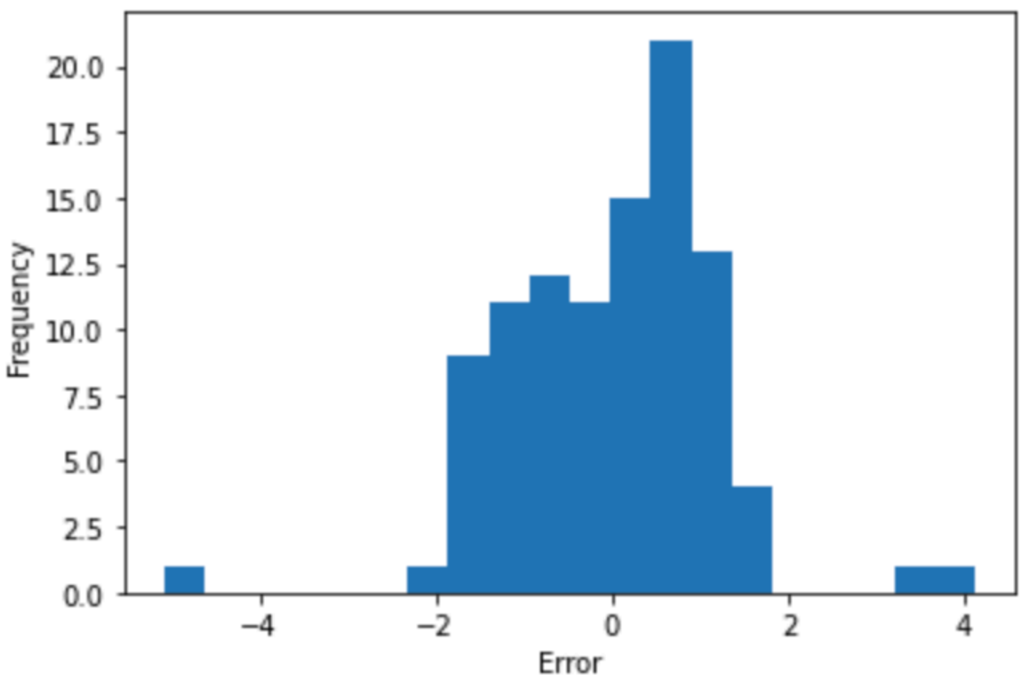
Problem 2
Fit the data in problem2.csv using OLS and calculate the error vector. Look at its distribution. How well does it fit the assumption of normally distributed errors?

Error vector:

```
0   -0.838485
1    0.835296
2    1.027428
3    1.319711
4   -0.152317
      ...
```

95 -1.590264
96 -1.694848
97 0.434878
98 0.402261
99 -0.922319

The distribution of the error vector is not normally distributed. It is not symmetric and a little skew to the left, with the mode around 1.



Fit the data using MLE given the assumption of normality. Then fit the MLE using the assumption of a T distribution of the errors. Which is the best fit? What are the fitted parameters of each and how do they compare? What does this tell us about the breaking of the normality assumption in regards to expected values in this case?

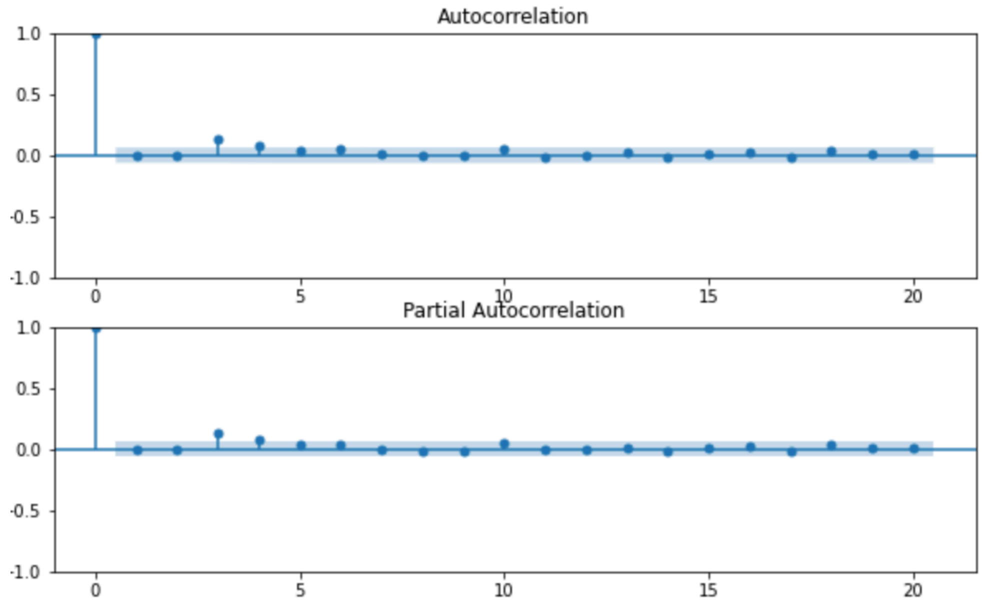
When fit the MLE with the assumption of a T distribution of the errors, it is the best fit. The reason is that ll is -164.98 when using MLE given the assumption of normality, while ll is -164.97 which is greater when using the assumption of a T distribution of the errors. Therefore, the one with the assumption of a T distribution of the errors has the best fit.

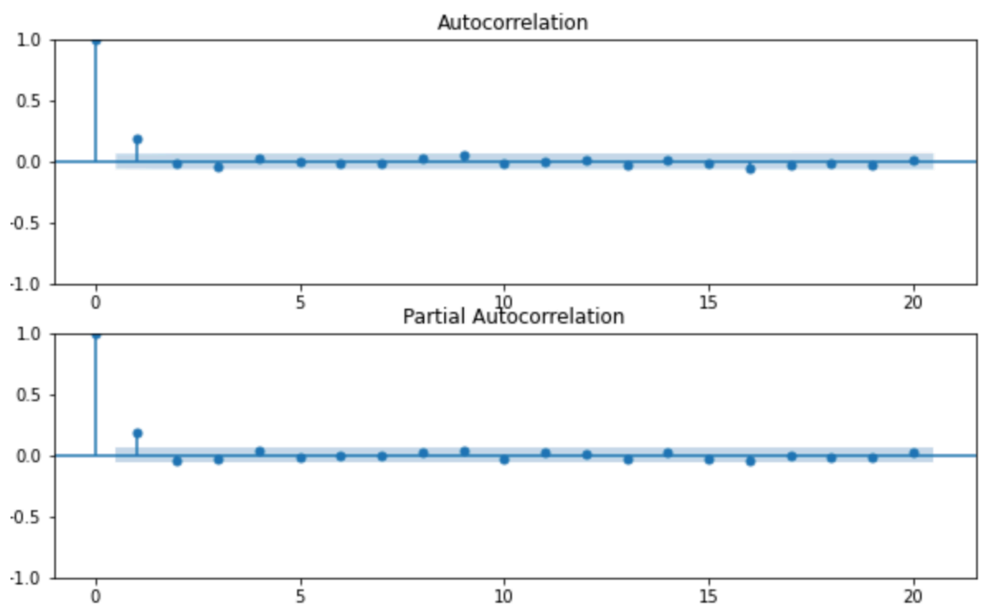
In general, the t-distribution should be considered as an alternative to the normal distribution if the data does not meet the assumption of normality since the the ll of two cases are pretty close.

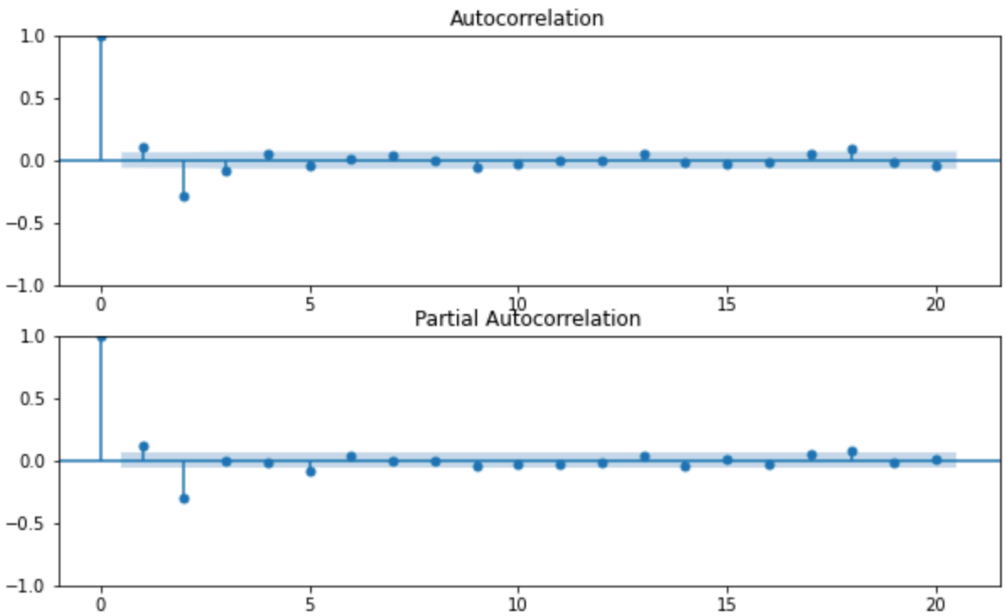
Problem3

Simulate AR(1) through AR(3) and MA(1) through MA(3) processes. Compare their ACF and PACF graphs. How do the graphs help us to identify the type and order of each process?

Here are the graphs of autocorrelation and partial autocorrelation of AR(1), AR(2), AR(3):

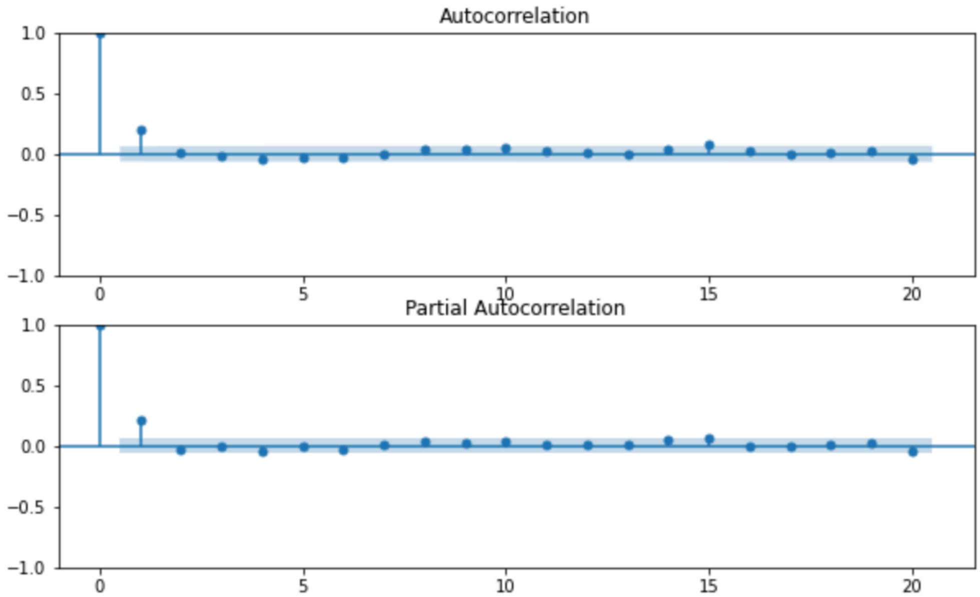


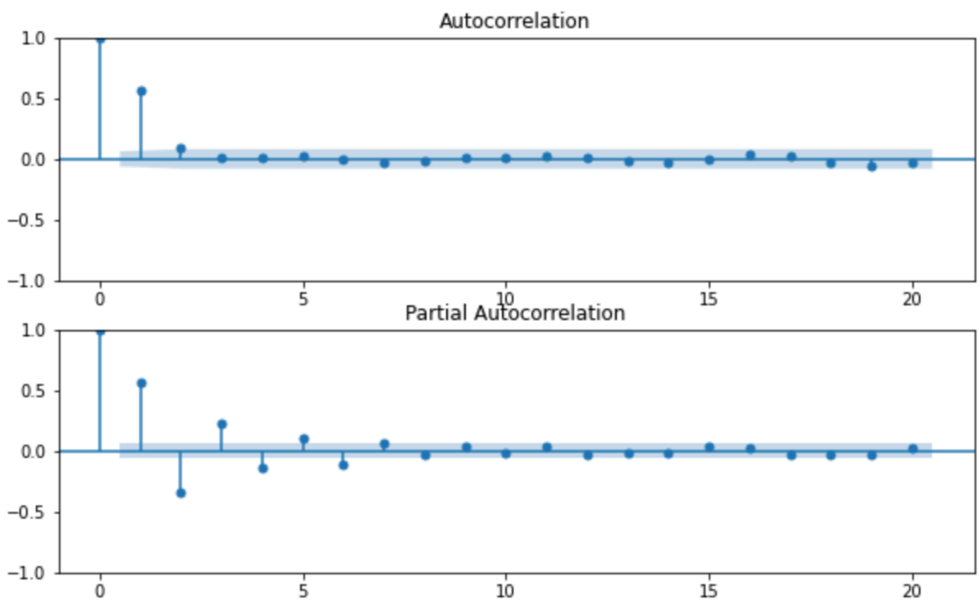


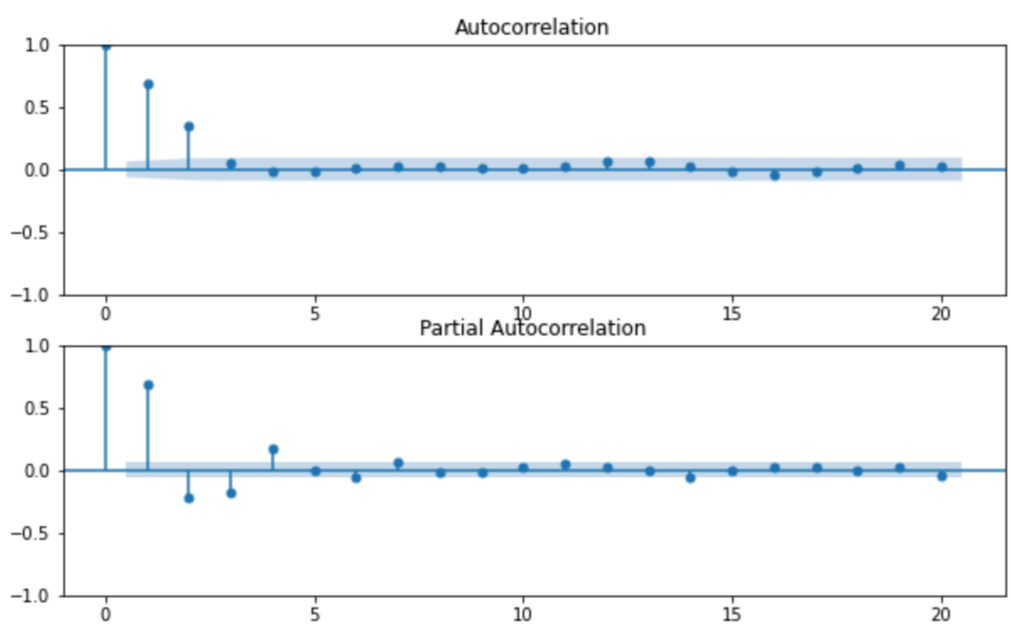


ACF has higher correlation at lower lag such as AR(1), while ACF has lower correlation at higher lag such and AR(3). Similarly, PACF has higher correlation at lower lag such as AR(1), while ACF has lower correlation at higher lag such and AR(3)

Here are the graphs of autocorrelation and partial autocorrelation of MA(1), MA(2), MA(3):


Autocorrelation
Partial Autocorrelation
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20


Autocorrelation
Partial Autocorrelation
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20


Autocorrelation
Partial Autocorrelation
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20

ACF has higher correlation at lower lag such as MA(1), while ACF has lower correlation at higher lag such and MA(3). Similarly, PACF has higher correlation at lower lag such as MA(1), while ACF has lower correlation at higher lag such and MA(3).