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.

```
Summary Stats:
Length:         1000
Missing Count:  0
Mean:           -0.569039
Minimum:        -1.805764
1st Quartile:   -1.066712
Median:         -0.732574
3rd Quartile:   -0.230604
Maximum:        3.588278
Type:           Float64
p-value - 0.0
Match the stats package test?: false
Summary Stats:
Length:         1000
Missing Count:  0
Mean:           -0.000189
Minimum:        -0.876372
1st Quartile:   -0.177615
Median:         0.005259
3rd Quartile:   0.162701
Maximum:        0.789665
Type:           Float64
p-value - 0.9803965263698999
Match the stats package test?: true
```

The above is the result from Julia. The first summary is the test of kurtosis function for bias, the second summary is the test of skewness function for bias. The sample size is 10 and is repeated 1000 times. The p-value for kurtosis function is 0 so we reject the null hypothesis that the kurtosis function is unbiased. It also doesn't match the stats package test. The p-value for skewness function is high. So, we cannot reject the null hypothesis that the skewness function is unbiased.

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?

Fit the data using MLE given the assumption of normality. The fit the MLE using the assumption of a T distribution of 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?

The distribution of the error vector doesn't fit the assumption of normally distributed errors well. The skewness is close to 0 but the kurtosis is 3.193 which shows that it has a fat fail.

```
                            OLS Regression Results
=======================================================================================
Dep. Variable:                      y   R-squared (uncentered):                   0.193
Model:                            OLS   Adj. R-squared (uncentered):              0.185
Method:                 Least Squares   F-statistic:                              23.69
Date:                Sat, 28 Jan 2023   Prob (F-statistic):                    4.28e-06
Time:                        01:42:44   Log-Likelihood:                         -160.49
No. Observations:                 100   AIC:                                      323.0
Df Residuals:                      99   BIC:                                      325.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1             0.6052      0.124      4.867      0.000       0.358       0.852
==============================================================================
Omnibus:                       14.146   Durbin-Watson:                   1.866
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               43.674
Skew:                          -0.267   Prob(JB):                     3.28e-10
Kurtosis:                       6.193   Cond. No.                         1.00
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
Skewness of error: -0.2672665855287956
```

```
Skewness of error: -0.2672665855287956
kurtosis of error: 3.1931010009568794
Normal betas: [0.11983616 0.60520482]
Normal s:1.1983941123422093
Normal ll:-159.9920966891624
Normal AIC:325.9841933783248
Normal BIC:333.79970393628906
T betas: [0.14261405 0.55757179]
T s:0.9712661373498599
T df:6.2765700023111615
T ll:-155.472970412474
T AIC:318.945940824948
T BIC:329.36662156890037
```

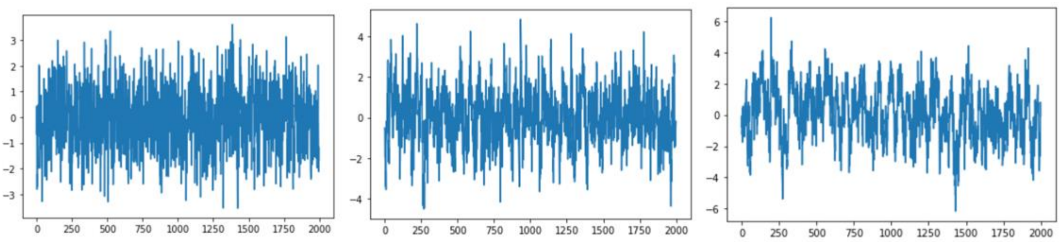
As we can see, the AIC, BIC and absolute value of log likelihood of using the assumption of a T distribution of the errors are smaller than that of normal distribution. So, using the T distribution is a better fit.

The breaking of the normality assumption is mostly showed by kurtosis. The AIC and BIC of using MLE are smaller than that of using OLS which shows that here MLE is a better fit.

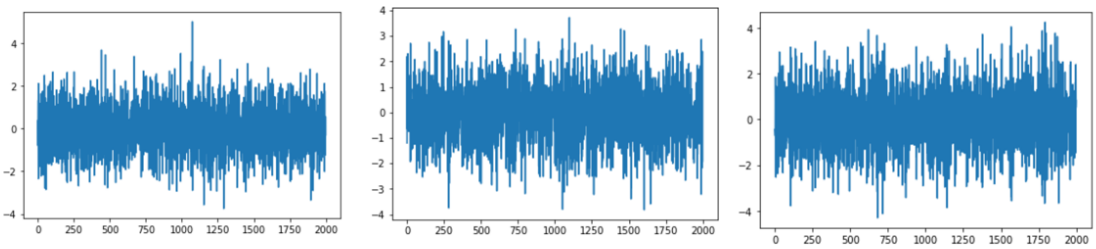
Problem 3:

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.

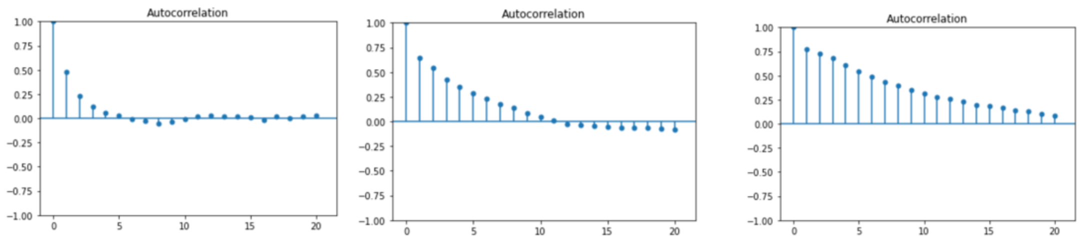
Simulation of AR(1) to AR(3):



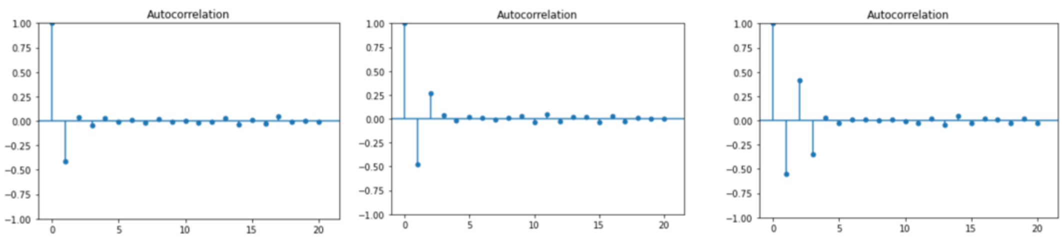
Simulation of MA(1) to MA(3):



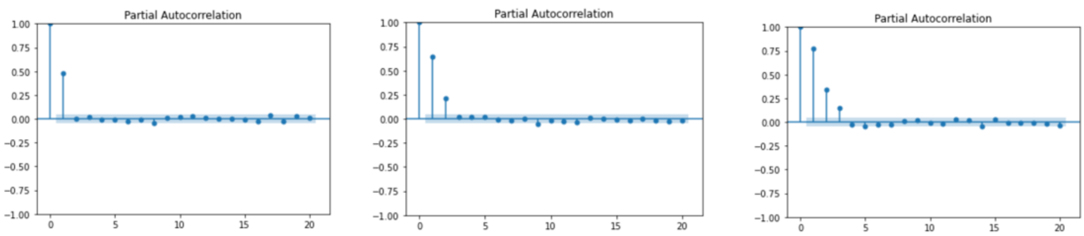
ACF graphs for AR(1) to AR(3):



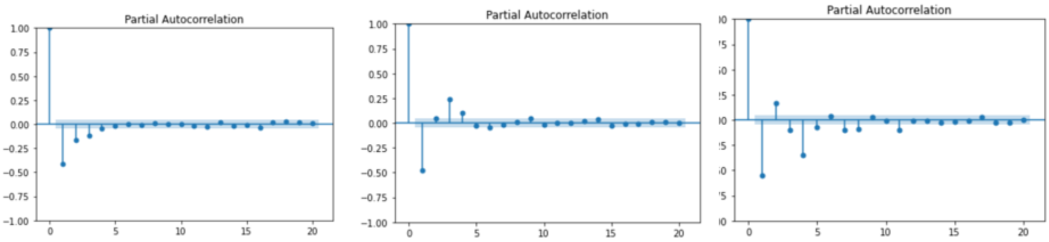
ACF graphs for MA(1) to MA(3):



PACF graphs for AR(1) to AR(3):



PACF graphs for MA(1) to MA(3):



The graphs show that as the order increases, ACF and PACF for AR and MA decays more smoothly and have more strong correlation. For PACF, many data are in the blue area which is statistically close to zero.