# Problem 1:

## a)

Mean: 1.0489703904839585
Variance: 5.4217934611998455
Skewness: 0.8806086425277359
Kurtosis: 26.12220078998972

## b)

Mean: 1.0489703904839585
Variance: 5.4217934611998455
Skewness: 0.8806086425277364
Kurtosis: 26.122200789989723

## c)

The p-values of four moments are:
Mean (np.mean): 0.17656961322179487
Variance (np.var): 0.007547098583956968
Skewness (scipy.stats.skewness): 0.7366114817578768
Excess Kurtosis (scipy.stats.kurtosis): 7.388381453532489e-06

Therefore, we can find that np.variance and scipy.stats.kurtosis is biased

# Problem 2

## a)

OLS: [-0.0874, 0.7753], Standard deviation of Error = 1.003756319417732
MLE: [-0.08738446, 0.7752741], Standard deviation of Error = 1.003756319417732

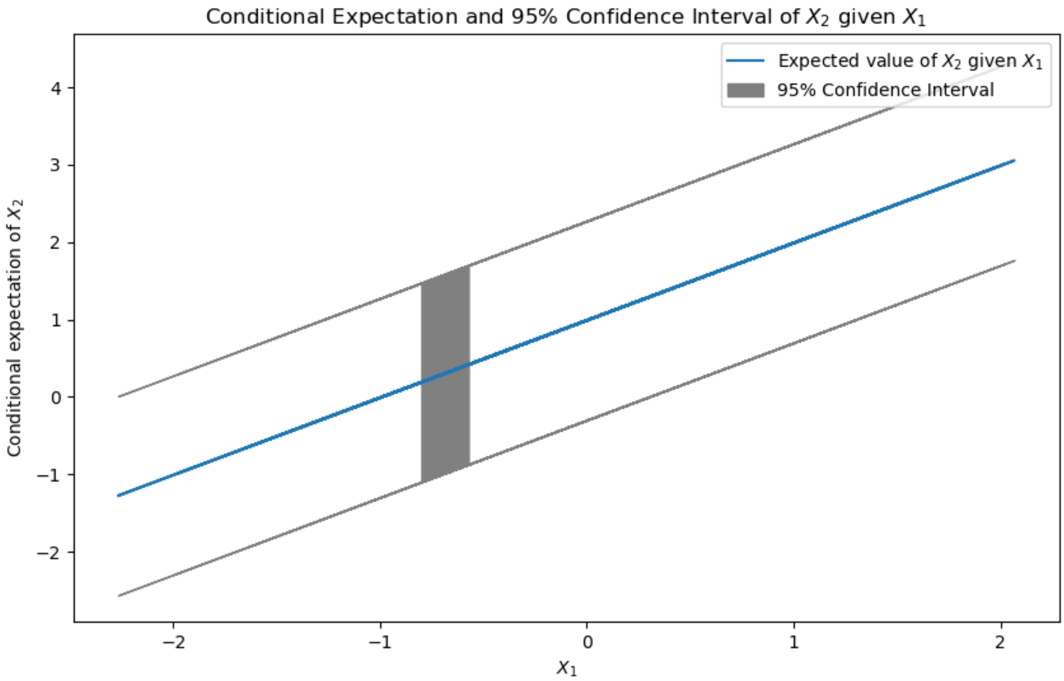
By comparing the result of OLS and MLE methods, we can find that their beta and the standard deviation of error (close to 1) is the same.

b)

MLE_N: [-0.08738446, 0.7752741], Adjusted R^2 = 0.3423018678252412
MLE_T: [-0.08884091, 0.76947098], Adjusted R^2= 0.3363196206776301

By comparing these two results, we can find that the result is almost the same, but MLE_N (Error under normal assumption) is slightly better.

c)



d)

The log likelihood function of $\beta, \sigma^2$:

$$ll(\beta, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n} \frac{(Y_i - X_i\beta)^2}{2\sigma^2}$$

$$\hat{\beta} = \operatorname{argmax}(ll(\beta, \sigma^2))$$

Set $\frac{\partial ll(\beta,\sigma^2)}{\partial \beta} = \frac{1}{\sigma^2}\sum_{i=1}^{n} X_i(Y_i - X_i\beta) = \frac{1}{\sigma^2}X^T(Y - X\beta) = 0$

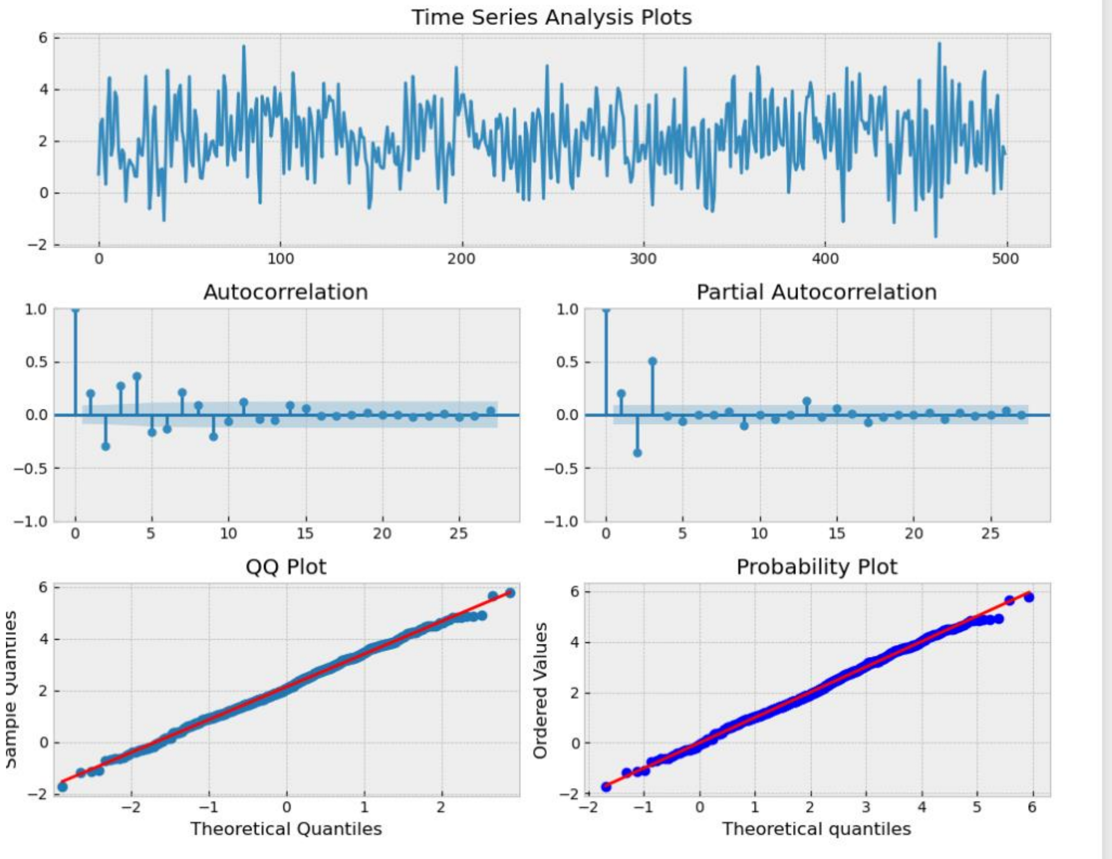
$$\Rightarrow \hat{\beta} = (X^TX)^{-1}X^TY$$

Similarly, set $\frac{\partial ll(\beta,\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(Y_i - X_i\beta)^2 = 0$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i\beta)^2$$

# Problem 3

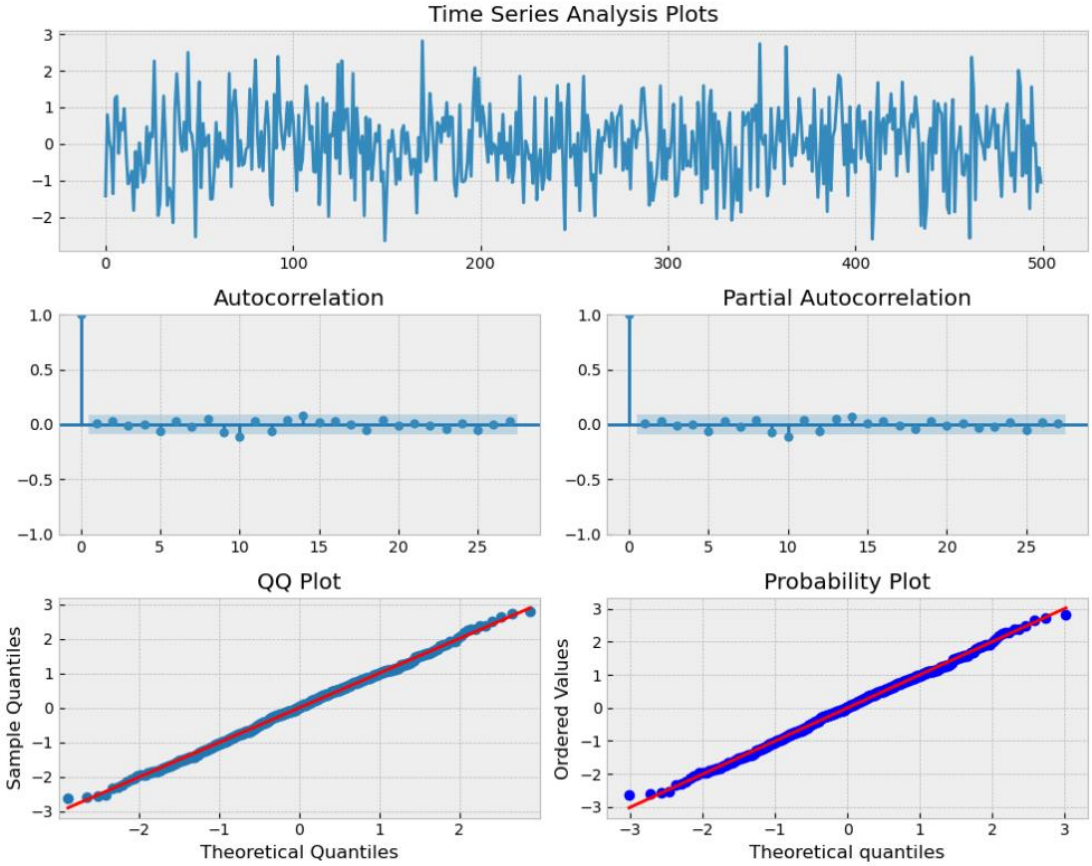
Firstly, we can judge the appropriate order by observing the image. By observing this graph, we can find that for ACF graph, the value of Autocorrelation gradually decreases to 0, and for PACF graph, the value of Partial Autocorrelation suddenly drops to 0, so we can infer that AR (3) model may be our good choice,




Besides, we can compare each model's AIC value to see which model is batter. By observing the following table, we can easily find that AR (3) has the lowest AIC value, which proves that AR (3) is the best model.

| | |
|---|---|
| ARIMA (1, 0, 0) | 1644.6555047688475 |
| ARIMA (2, 0, 0) | 1581.079265904977 |
| ARIMA (3, 0, 0) | 1436.6598066945883 |
| ARIMA (0, 0, 1) | 1567.4036263707874 |
| ARIMA (0, 0, 2) | 1537.941206380739 |
| ARIMA (0, 0, 3) | 1536.8677087350302 |

Finally, we can draw the graph of the residuals of AR (3) model to see if the residuals look like the white noise, and from the following graph, it can be easily find that the residuals is white noise because QQ plot and Probability plot shows it is normally distributed and ACF and PACF plot shows that residuals don't have any Autocorrelation with past values.


Time Series Analysis Plots
Autocorrelation
Partial Autocorrelation
QQ Plot
Sample Quantiles
Theoretical Quantiles
Probability Plot
Ordered Values
Theoretical quantiles