Week 2 Homework for Fintech545

Yue Yu (yy404)

Question 1

a.

b. I used excel to type formulas to calculate 4 moments.

Moment	Normalized formula	Unnormalized formula	Biased formula	python
Mean	1.04897	1.04897	1.04897	1.04897
Variance	5.42722	5.42722	5.421793	5.421793
Skewness	0.88193	11.15068	11.11725	11.11725
Kurtosis	23.07013	767.88858	767.88414	767.88414

c. Biased and unnormalized. Compared to the skewness and kurtosis of biased formula and unnormalized formula, python package scipy is biased and unnormalized.

Question 2

a. Regression – OLS

In OLS regression, the coefficient (beta) for the variable x is 0.7753, for the constant is -0.0874, while the standard deviation is 1.004.

```
Standard Deviation of OLS Error: 1.003756319417732
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.346
Model:                            OLS   Adj. R-squared:                  0.342
Method:                 Least Squares   F-statistic:                     104.6
Date:                Thu, 25 Jan 2024   Prob (F-statistic):           5.59e-20
Time:                        14:42:58   Log-Likelihood:                -284.54
No. Observations:                 200   AIC:                             573.1
Df Residuals:                     198   BIC:                             579.7
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
constant      -0.0874      0.071     -1.222      0.223      -0.228       0.054
x              0.7753      0.076     10.226      0.000       0.626       0.925
==============================================================================
Omnibus:                       11.922   Durbin-Watson:                   2.023
Prob(Omnibus):                  0.003   Jarque-Bera (JB):               16.685
Skew:                           0.387   Prob(JB):                     0.000238
Kurtosis:                       4.184   Cond. No.                         1.09
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

In MLE following the assumption of normality, MLE estimated mean is -0.038, standard deviation is 1.241 and beta is 0.7753.

```
  message: Optimization terminated successfully.
  success: True
   status: 0
      fun: 284.5375630544279
        x: [-8.738e-02  7.753e-01  1.004e+00]
      nit: 14
      jac: [ 0.000e+00  3.815e-06 -3.815e-06]
 hess_inv: [[ 5.126e-03 -4.949e-04  6.627e-07]
            [-4.949e-04  5.900e-03  1.627e-06]
            [ 6.627e-07  1.627e-06  2.592e-03]]
     nfev: 72
     njev: 18
```

Obviously, the OLS model exhibits a better fit as its standard deviation is smaller, implying that the model's predicted values are closer to the actual observations. The MLE model shows a relatively poorer fit since its standard deviation is larger, indicating a larger discrepancy between the predicted values and the actual observations. We can see that the beta of MLE is the same as the OLS solution.

b. In MLE under the assumption of t distribution, MLE estimate of mean is 0.003, standard deviation is 0.930 and degrees of freedom is 7.051.

	MLE_Norm	MLE_t
Log Likelihood	-326.942	-597.126
AIC	657.88	1200.252
BIC	664.48	1210.147

Comparing the MLE_Norm and MLE_t models based on Log Likelihood, AIC, and BIC values, we can draw the following conclusions:

1. Log Likelihood:

- The Log Likelihood for the MLE_t model is -597.126, while for the MLE_Norm model, it is -326.942.
- Since a higher Log Likelihood indicates a better fit of the model to the data, the MLE_Norm model performs better in this aspect.

2. AIC (Akaike Information Criterion):

- The AIC for the MLE_t model is 1200.252, whereas for the MLE_Norm model, it is 657.88.
- AIC considers both model fit and complexity, and a lower AIC value suggests a relatively better model. Therefore, in terms of AIC, the MLE_Norm model performs better.

3. BIC (Bayesian Information Criterion):

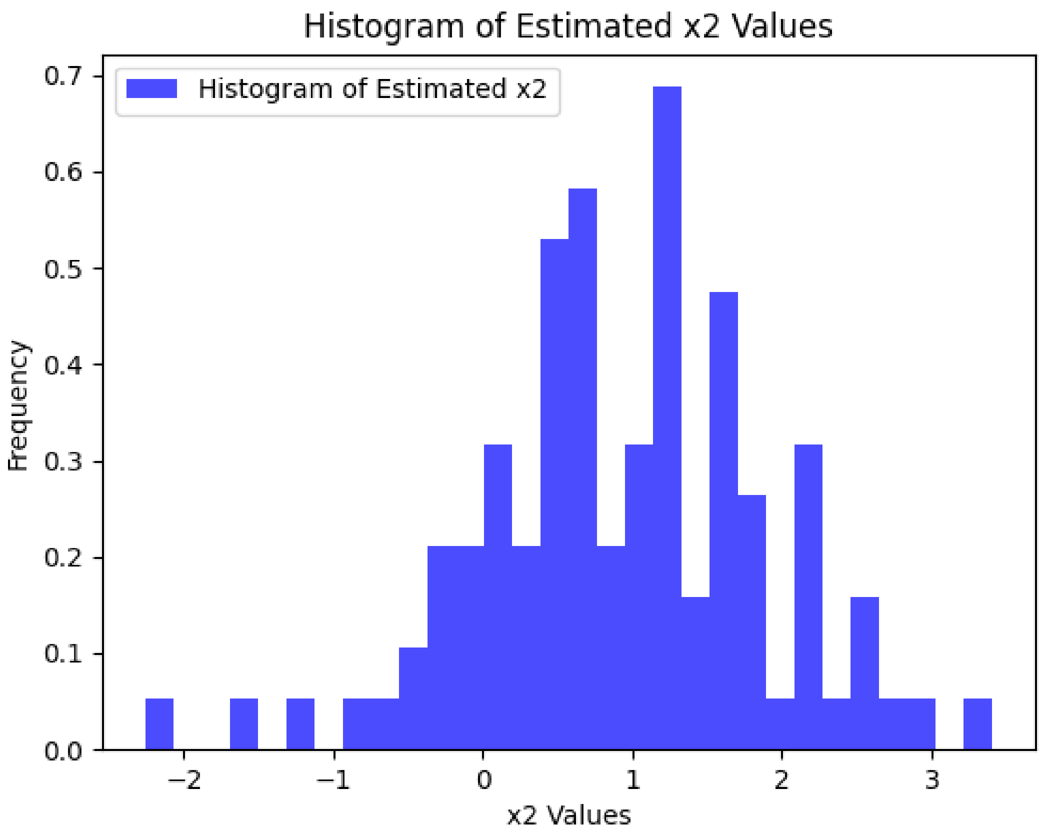
- The BIC for the MLE_t model is 1210.147, while for the MLE_Norm model, it is 664.48.
- Similar to AIC, BIC considers both fit and complexity, and a lower BIC value indicates a relatively better model. In terms of BIC, the MLE_Norm model also performs better.

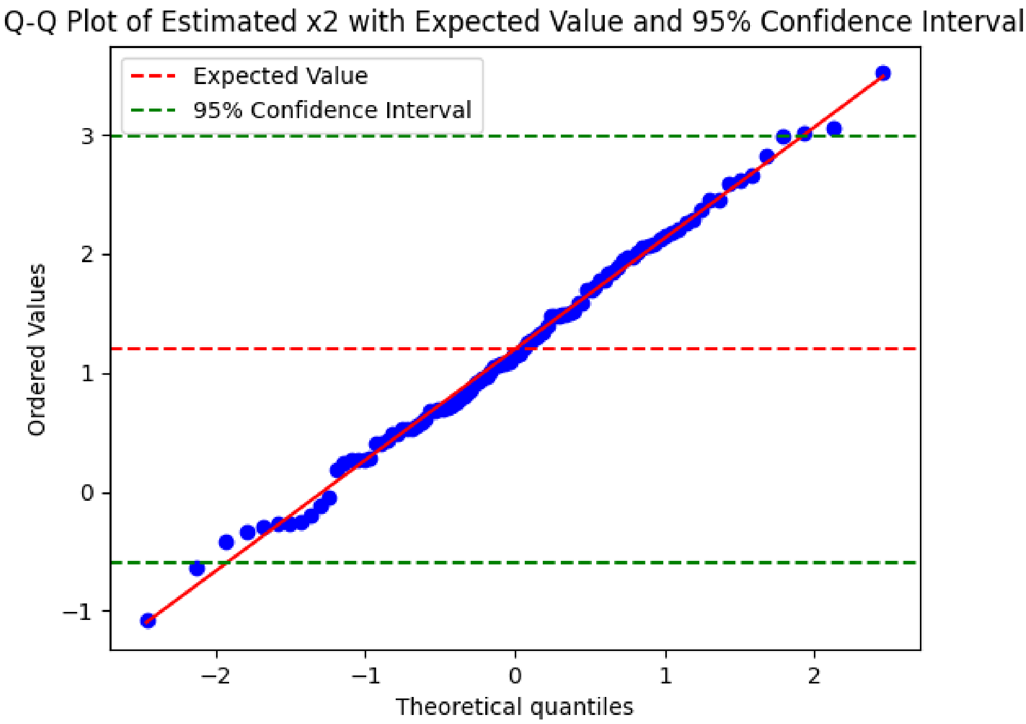
Taking into account Log Likelihood, AIC, and BIC collectively, the MLE_Norm model outperforms the MLE_t model in this comparison, exhibiting both better fit to the data and a simpler model.

c. In MLE following the multivariate normal distribution, estimated mean vector is $\begin{bmatrix} 0.0010227 & 0.99024382 \end{bmatrix}$ and estimated covariance matrix is

$$\begin{bmatrix} 1.06977464 & 0.53068455 \\ 0.53068455 & 0.96147329 \end{bmatrix}$$

In order to estimate X_2 from observed X_1, I created a multivariate normal distribution object for X_2 given X_1 at the very beginning, then I used Random Value Sample to get the X_2 value. After that, I calculate the mean and confidence interval for the estimated X_2, which is [0.99024382] and [[0.96147329]] respectively. After plotting histogram graph of estimated x2, I guessed it follows normal distribution.Then, plot Q-Q plot with the expected value and 95% confidence interval to normal distribution. From the Q-Q plot, it is evident that the distribution of X2 is well-fitted to a normal distribution.





d. Please see the detailed prove in last page.

$$β = (X^T X)^{-1} X^T y$$

$$σ^2 = \frac{1}{N} \sum_{i=1}^{N} ((y_i - x_i\, β)\text{^}2)$$

Question 3

According to the models built, log likelihood, AIC and BIC can be used to analyze their performance:

	AR(1)	AR(2)	AR(3)	MA(1)	MA(2)	MA(3)
Log Likelihood	-819.328	-786.540	-713.330	-780.702	-764.971	-763.434
AIC	1644.656	1581.079	1436.660	1567.404	1537.941	1536.868
BIC	1657.299	1597.938	1457.733	1580.047	1554.800	1557.941

Analyzing the provided values for different AR and MA models:

1. Log Likelihood: A higher log-likelihood value indicates better model fit. In this case, AR(3) has the highest log-likelihood (-713.330), suggesting it fits the data best among the specified models.

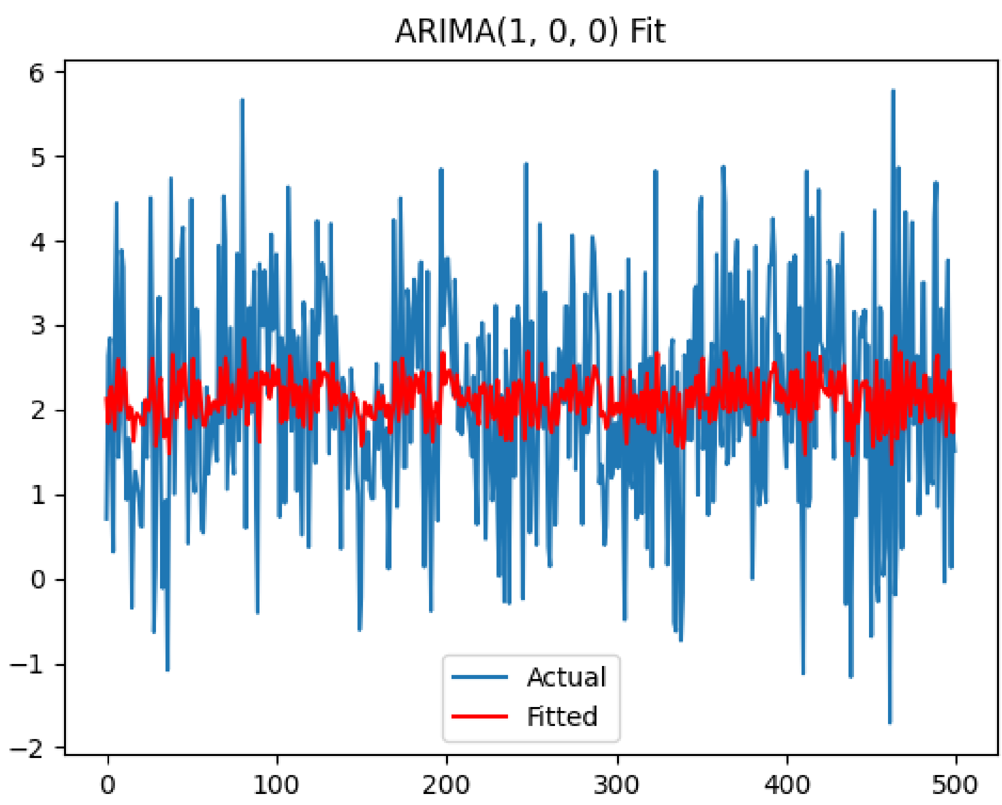
2. AIC (Akaike Information Criterion): Lower AIC values indicate better model fit, balancing goodness of fit and model complexity. AR(3) has the lowest AIC (1436.660), suggesting it provides a better trade-off between fit and complexity compared to other models.

3. BIC (Bayesian Information Criterion): Similar to AIC, lower BIC values indicate better model fit with a penalty for model complexity. AR(3) also has the lowest BIC (1457.733), supporting its better fit relative to the other models.

In summary, based on the log-likelihood, AIC, and BIC values, the AR(3) model appears to provide the best fit among the specified AR and MA models for the given dataset.

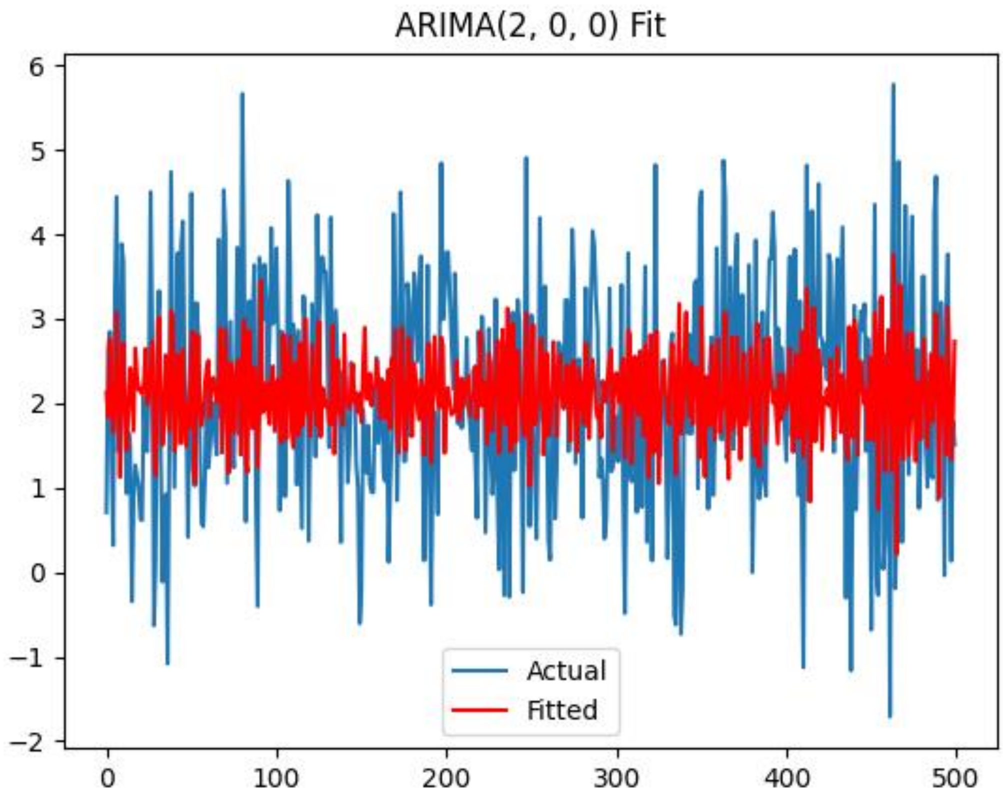
AR(1)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      x   No. Observations:                  500
Model:                 ARIMA(1, 0, 0)   Log Likelihood                -819.328
Date:                Thu, 25 Jan 2024   AIC                           1644.656
Time:                        16:29:59   BIC                           1657.299
Sample:                             0   HQIC                          1649.617
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1258      0.070     30.473      0.000       1.989       2.263
ar.L1          0.2019      0.045      4.512      0.000       0.114       0.290
sigma2         1.5517      0.105     14.743      0.000       1.345       1.758
===================================================================================
Ljung-Box (L1) (Q):                   2.51   Jarque-Bera (JB):                 1.42
Prob(Q):                              0.11   Prob(JB):                         0.49
Heteroskedasticity (H):               1.37   Skew:                            -0.00
Prob(H) (two-sided):                  0.04   Kurtosis:                         2.74
===================================================================================


Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
AR(1) MSE: 2.188192947470165
```



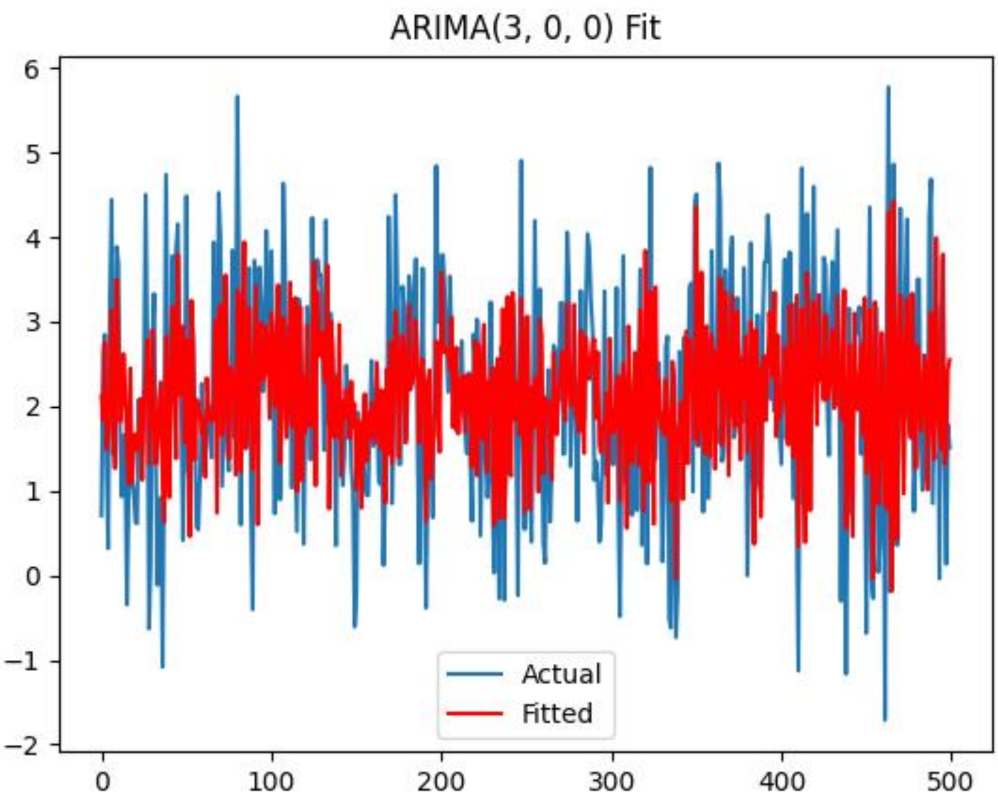
AR(2)

```
                                SARIMAX Results
==============================================================================
Dep. Variable:                      x   No. Observations:                  500
Model:                 ARIMA(2, 0, 0)   Log Likelihood                -786.540
Date:                Thu, 25 Jan 2024   AIC                           1581.079
Time:                        16:30:25   BIC                           1597.938
Sample:                             0   HQIC                          1587.694
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1270      0.049     43.663      0.000       2.032       2.222
ar.L1          0.2732      0.042      6.486      0.000       0.191       0.356
ar.L2         -0.3505      0.043     -8.068      0.000      -0.436      -0.265
sigma2         1.3603      0.094     14.455      0.000       1.176       1.545
===================================================================================
Ljung-Box (L1) (Q):                  15.51   Jarque-Bera (JB):                 3.12
Prob(Q):                              0.00   Prob(JB):                         0.21
Heteroskedasticity (H):               1.20   Skew:                            -0.11
Prob(H) (two-sided):                  0.24   Kurtosis:                         2.68
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
AR(2) MSE: 2.18252967499774
```



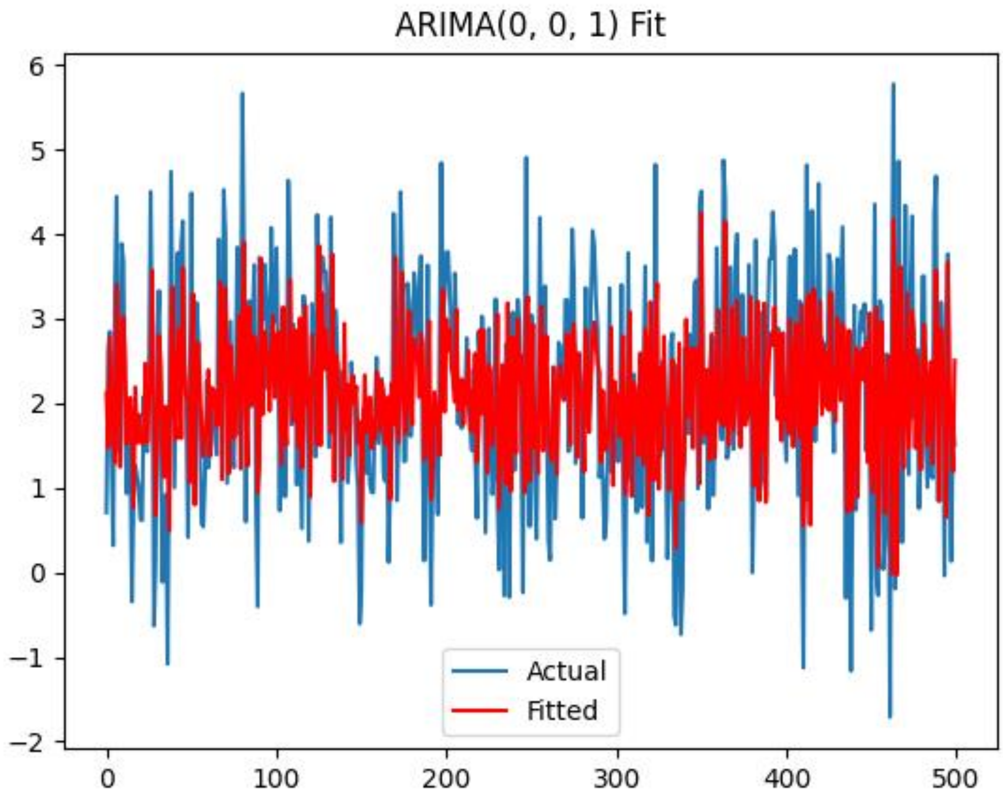
AR(3)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      x   No. Observations:                  500
Model:                 ARIMA(3, 0, 0)   Log Likelihood                -713.330
Date:                Thu, 25 Jan 2024   AIC                           1436.660
Time:                        16:30:41   BIC                           1457.733
Sample:                             0   HQIC                          1444.929
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1209      0.085     24.990      0.000       1.955       2.287
ar.L1          0.4515      0.040     11.179      0.000       0.372       0.531
ar.L2         -0.4887      0.037    -13.104      0.000      -0.562      -0.416
ar.L3          0.5047      0.040     12.769      0.000       0.427       0.582
sigma2         1.0132      0.068     14.939      0.000       0.880       1.146
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):                 0.84
Prob(Q):                              0.90   Prob(JB):                         0.66
Heteroskedasticity (H):               1.04   Skew:                            -0.03
Prob(H) (two-sided):                  0.81   Kurtosis:                         2.81
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
AR(3) MSE: 2.1882549284358204
```



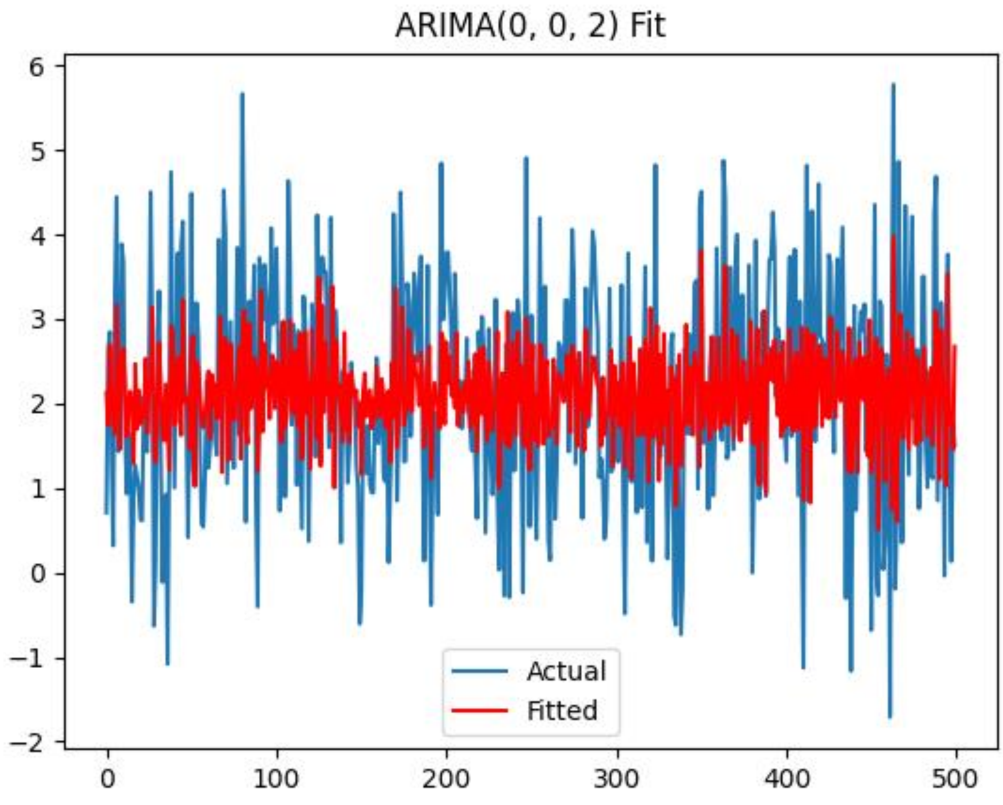
MA(1)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      x   No. Observations:                  500
Model:                 ARIMA(0, 0, 1)   Log Likelihood                -780.702
Date:                Thu, 25 Jan 2024   AIC                           1567.404
Time:                        16:31:17   BIC                           1580.047
Sample:                             0   HQIC                          1572.365
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1236      0.085     25.028      0.000       1.957       2.290
ma.L1          0.6434      0.034     18.847      0.000       0.577       0.710
sigma2         1.3282      0.090     14.782      0.000       1.152       1.504
===================================================================================
Ljung-Box (L1) (Q):                  11.73   Jarque-Bera (JB):                 1.18
Prob(Q):                              0.00   Prob(JB):                         0.55
Heteroskedasticity (H):               1.39   Skew:                            -0.02
Prob(H) (two-sided):                  0.04   Kurtosis:                         2.77
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```



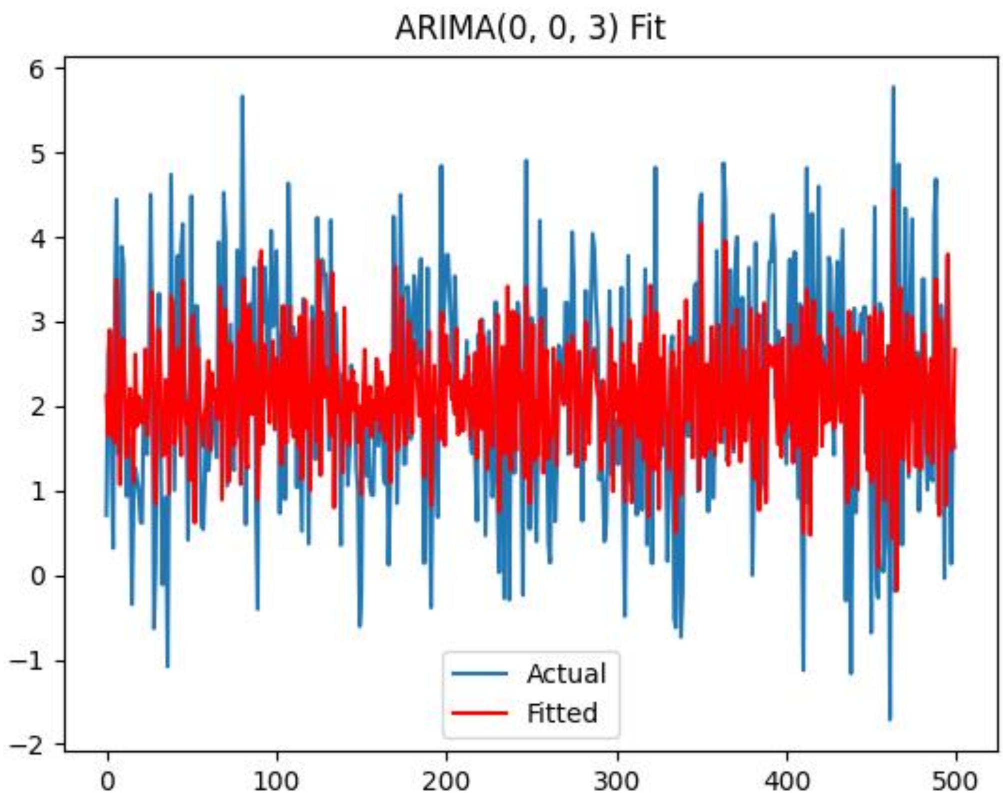
MA(2)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      x   No. Observations:                  500
Model:                 ARIMA(0, 0, 2)   Log Likelihood                -764.971
Date:                Thu, 25 Jan 2024   AIC                           1537.941
Time:                        16:31:40   BIC                           1554.800
Sample:                             0   HQIC                          1544.556
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1255      0.060     35.199      0.000       2.007       2.244
ma.L1          0.4344      0.044      9.775      0.000       0.347       0.522
ma.L2         -0.2306      0.047     -4.949      0.000      -0.322      -0.139
sigma2         1.2473      0.086     14.558      0.000       1.079       1.415
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):                 1.67
Prob(Q):                              0.88   Prob(JB):                         0.43
Heteroskedasticity (H):               1.28   Skew:                            -0.03
Prob(H) (two-sided):                  0.11   Kurtosis:                         2.72
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```



MA(3)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      x   No. Observations:                  500
Model:                 ARIMA(0, 0, 3)   Log Likelihood                -763.434
Date:                Thu, 25 Jan 2024   AIC                           1536.868
Time:                        16:31:51   BIC                           1557.941
Sample:                             0   HQIC                          1545.137
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1259      0.059     35.880      0.000       2.010       2.242
ma.L1          0.5582      0.045     12.333      0.000       0.469       0.647
ma.L2         -0.2286      0.053     -4.308      0.000      -0.333      -0.125
ma.L3         -0.1531      0.048     -3.216      0.001      -0.246      -0.060
sigma2         1.2394      0.085     14.592      0.000       1.073       1.406
===================================================================================
Ljung-Box (L1) (Q):                   1.60   Jarque-Bera (JB):                 1.75
Prob(Q):                              0.21   Prob(JB):                         0.42
Heteroskedasticity (H):               1.25   Skew:                            -0.06
Prob(H) (two-sided):                  0.15   Kurtosis:                         2.73
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```



$$\max_{\beta,\sigma^2} l(\beta, \sigma^2; y, X)$$

derivative $\to \nabla_\beta l(\beta, \sigma^2; y, X) = 0$

$$\frac{\partial}{\partial \sigma^2} l(\beta, \sigma^2; Y, X) = 0$$

$$\nabla_\beta l(\beta, \sigma^2; y, X)$$

$$= \nabla_\beta \left(-\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2\right)$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T (y_i - x_i\beta)$$

$$= \frac{1}{\sigma^2}\left(\sum_{i=1}^{N} x_i^T y_i - \sum_{i=1}^{N} x_i^T x_i \beta\right)$$

when $\sum_{i=1}^{N} x_i^T y_i = \sum_{i=1}^{N} x_i^T x_i \beta$, it equal to zero

$$\beta = \left(\sum_{i=1}^{N} x_i^T x_i\right)^{-1} \sum_{i=1}^{N} x_i^T y_i = (X^T X)^{-1} X^T y$$

$$\frac{\partial}{\partial \sigma^2} l(\beta, \sigma^2; y, X)$$

$$= \frac{\partial}{\partial \sigma^2}\left(-\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2\right)$$

$$= \frac{1}{2\sigma^2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2 - N\right]$$

when $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i\beta)^2$

$$\beta = (X^T X)^{-1} X^T y \qquad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i\beta)^2$$