Week02_Report
Wanjing Xu

Question 1
a.
n_x = 1000

1) $\hat{\mu}_1 = E[X] = \frac{1}{n}\sum_{i}^{n} x_i$

sum of x = 1048.970390
the mean of x = 1 / 1000 * 1048.970390 = 1.0490

2) $\hat{\mu}_2 = E[(X - \hat{\mu}_1)^2] = \frac{1}{n-1}\sum_{i}^{n}(x_i - \hat{\mu}_1)^2$

sum of (x-mean)^2 = 5421.793461
the variance of x = 1 / 999 * 5421.793461 = 5.4272

3) $\hat{\mu}_3 = E[(\frac{X-\mu}{\sigma})^3] = \frac{u_3}{\sigma^3}$

sum of (x-mean)^3 = 11117.25065592259
skewness of x = sum of (x-mean)^3/n/sqrt(sigma_x^2) = 11117.25065592259 / 1000 / sqrt(5.421793461199845^3) = 0.8806

4) $\hat{\mu}_4 = E[(\frac{X-\mu}{\sigma})^4] - 3 = \frac{u_4}{\sigma^4} - 3$

sum of (x-mean)^4 = 767884.148133907
kurtosis of x = sum of (x-mean)^4/n/sigma_x^2 = 465.9228484697776 / 1000 / 5.421793461199845^2 = 26.1222
excess kurtosis of x = kurtosis of x - 3 = 23.1222

b. I chose python pandas and numpy statistical packages. I calculate the difference between x and its mean and the power for the difference from 2 to 4. The second part involves four moments calculated by normalized formula and by pandas.

```
            x    x-mean  (x-mean)^2  (x-mean)^3  (x-mean)^4
0    0.922177 -0.126794    0.016077   -0.002038    0.000258
1    1.791781  0.742810    0.551767    0.409858    0.304447
2    0.397551 -0.651419    0.424347   -0.276428    0.180071
3    3.564269  2.515298    6.326725   15.913600   40.027451
4   -0.691414 -1.740385    3.028938   -5.271517    9.174467
..        ...       ...         ...         ...         ...
995 -2.913057 -3.962027   15.697660  -62.194556  246.416523
996  2.845173  1.796203    3.226344    5.795168   10.409296
997  1.663335  0.614365    0.377444    0.231888    0.142464
998  2.344441  1.295470    1.678243    2.174114    2.816499
999  1.164175  0.115204    0.013272    0.001529    0.000176

[1000 rows x 5 columns]
mean_x is 1.0489703904839585
var_x is 5.427220681681682
skewness of x is 0.8806086425277363
kurtosis_x is 26.122200789989733
ex_kurtosis_x is 23.122200789989733
pandas mean is 1.0489703904839585
pandas variance is 5.421793461199845
pandas skewness is 0.8806086425277365
pandas kurtosis is 26.122200789989733
```

c.

My statistical package is unbiased. Compared four moments calculated by normalized formula and by pandas, the results are the same.

Question 2

a.

OLS beta is 0.7753; OLS intercept is -0.0874; OLS standard deviation of residuals is 1.0038

MLE beta is 0.7753; MLE intercept is -0.0874; MLE standard deviation of residuals is 1.0038

Beta and standard deviation of residuals are the same under the OLS and MLE given the assumption of normality.

```
                            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:                        15:03:27   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]
------------------------------------------------------------------------------
const         -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.
```

```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  200
Model:                            GLM   Df Residuals:                      198
Model Family:                Gaussian   Df Model:                            1
Link Function:               identity   Scale:                          1.0177
Method:                          IRLS   Log-Likelihood:                -284.54
Date:                Thu, 25 Jan 2024   Deviance:                       201.51
Time:                        15:11:56   Pearson chi2:                     202.
No. Iterations:                     3   Pseudo R-squ. (CS):             0.4072
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0874      0.071     -1.222      0.222      -0.228       0.053
x              0.7753      0.076     10.226      0.000       0.627       0.924
==============================================================================
```

b.

Estimated Parameters - t distribution:

log_likelihood_t is 281.2934031796488

mle_beta: 0.675

mle_intercept: -0.0973

df: 7.1598

sigma: 0.8551

Estimated Parameters - under normality:

log_likelihood_n is 284.53756305442874

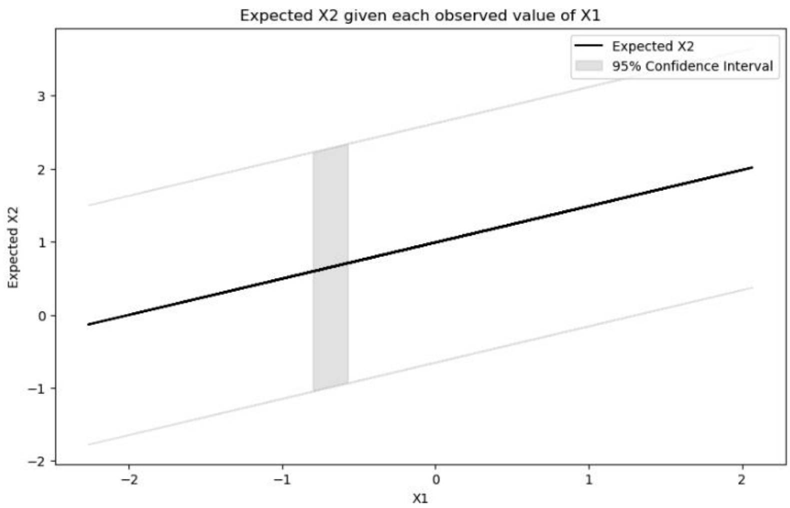
mle_beta: 0.7753

mle_intercept: -0.0874
sigma: 1.0038

AIC for t-distribution model: 570.5868
AIC for normal distribution model: 575.0751

The fitted parameters among MLE under normality assumption is slightly different than under T distribution assumption. An additional parameter under T distribution assumption allows greater adaptability for modeling. Sigma is smaller under T distribution assumption, suggesting that the t distribution model may have less variability in the errors than the normal distribution model. Compared with log likelihood values, MLE under T distribution assumption has a lower value, which shows a better fit. Similarly, the t-distribution model has a smaller AIC (570.5868), which has a better fit compared to the normal distribution model (575.0751).

c.
X = [X1, X2] follows the multivariate normal distribution.
When we plot the conditional mean of X2, X2 follows the normal distribution given the observed value of X1.



d.

Q2 d. $\varepsilon = Y - X\hat{\beta}$

Assume $\varepsilon \sim N(0, \sigma^2 I_n)$

$$f(y|\beta,\sigma^2) = \prod_{i=1}^{n} f(y_i|\beta,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu_i)^2}{2\sigma^2}} \qquad \mu_i = E(y_i) = x_i'\beta$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2}$$

$$\ln f(y|\beta,\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta)$$

① $\frac{\partial}{\partial\sigma^2} \ln f(y|\beta,\sigma^2) = 0$

$$0 - \frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(y-X\beta)'(y-X\beta) = 0$$

$$\frac{1}{\cancel{2}(\sigma^2)^{\cancel{2}}}(y-X\beta)'(y-X\beta) = \frac{n}{\cancel{2}\sigma^{\cancel{2}}}$$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n}(y-X\beta)'(y-X\beta)$$

② $\frac{\partial}{\partial\beta} \ln f(y|\beta,\sigma^2) = 0$

$$(y-X\beta)'(y-X\beta) = y'y - y'X\beta - \beta'X'y + \beta'X'X\beta$$

$$\frac{\partial(y'y)}{\partial\beta} = 0 \qquad \frac{\partial(-y'X\beta)}{\partial\beta} = (-y'X)' = -X'y$$

$$\frac{\partial(\beta'X'y)}{\partial\beta} = X'y \qquad \frac{\partial(\beta'X'X\beta)}{\partial\beta} = X'X\beta + X'X\beta$$

$$0 - 0 + 0 - X'y - X'y + X'X\beta + X'X\beta = 0$$

$$-2X'y + 2X'X\beta = 0$$

$$-X'y + X'X\beta = 0$$

$$X'X\beta = X'y$$

$$\cancel{(X'X)^{-1}(X'X)}\beta = (X'X)^{-1}X'y$$

$$\hat{\beta}_{MLE} = (X'X)^{-1}X'y$$

Question 3

Among AR models, the order 3 has the smallest AIC (1436.660) which means a better fit. Among MA models, the order 3 has the smallest AIC (1536.868) which means a better fit. Compared with AR(3), AR(3) has the smaller AIC (1436.660), which is the best of fit.

AR(1) p=1 d=0 q=0

```
                               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:                        22:50:54   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(2) p=2 d=0 q=0

```
                               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:                        22:52:05   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(3) p=3 d=0 q=0

```
                               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:                        22:52:54   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).
```

MA(1) p=0 d=0 q=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:                        22:53:49   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) p=0 d=0 q=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:                        22:53:49   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) p=0 d=0 q=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:                        22:55:35   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).
```