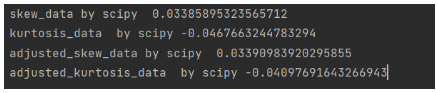
Problem 1: Remember from last week we discussed that skewness and kurtosis functions in statistical packages are often biased. Is your function biased? Prove or disprove your hypothesis.

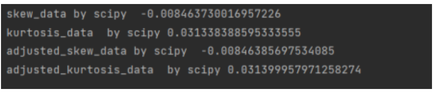
Answer:
I have randomly selected 1,000, 10,000 and 100,000 data in Python NumPy that fit a normal distribution. Theoretically, all three datasets should have a skewness of 0 and an excess kurtosis of 0. In practice, however, there is some error based on the test in biased or not in biased, which I attribute to the number of observations. The more data there is, the more it conforms to a normal distribution, as in the picture I've attached. Therefore, I think statistical packages tend to be biased because we usually have no control over the amount of data we can collect.

Test results:



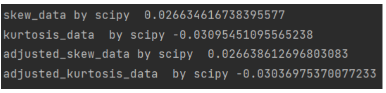
```
skew_data by scipy  0.03385895323565712
kurtosis_data  by scipy -0.0467663244783294
adjusted_skew_data by scipy  0.03390983920295855
adjusted_kurtosis_data  by scipy -0.04097691643266943
```

1000 observations



```
skew_data by scipy  -0.008463730016957226
kurtosis_data  by scipy 0.031338388595333555
adjusted_skew_data by scipy  -0.00846385697534085
adjusted_kurtosis_data  by scipy 0.031399957971258274
```

10,000 observations

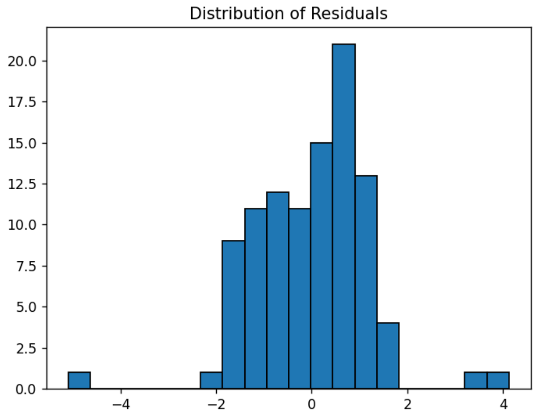


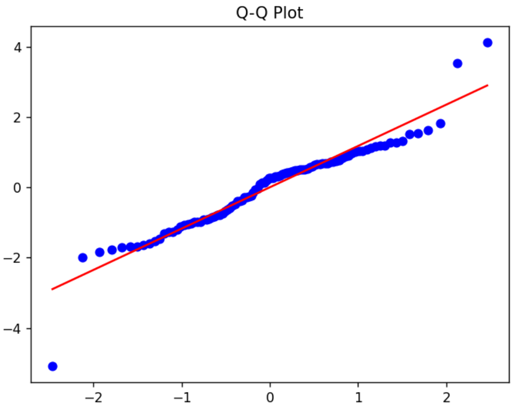
```
skew_data by scipy  0.026634616738395577
kurtosis_data  by scipy -0.03095451095565238
adjusted_skew_data by scipy  0.026638612696803083
adjusted_kurtosis_data  by scipy -0.03036975370077233
```

100,000 observations

Problem 2.1: 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?

Answer: Please run the Problem 2.1 code for the error vector. Here is the error vector distribution. And by Shapiro-Wilk Test, I concluded that at the 95% confidence level, errors do not follow a normal distribution.





Problem 2.2: 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?

Answerr: According to AIC and BICs, t distribution performs better.

Problem 2.3: 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 regard to expected values in this case?

Answer:

Given the assumption of normal distribution, MLE parameters are: [0.60520486 0.11983621 1.19839409]

Given the assumption of t-distribution, MLE parameters are: [0.55757171 0.1426142 6.27654811 0.97126583]

We will compare the two models by the last term (scale, which is the standard deviation).

The t-distribution will be relatively accurate in small sample statistics.

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

AR/MA (1) Model:

ACF (Autocorrelation Function): In the ACF plot of an AR(1) process, there will be a significant spike at lag 1, and the autocorrelations for all other lags will be close to zero.

PACF (Partial Autocorrelation Function): The PACF plot will have a significant spike at lag 1, and all other lags will be close to zero. There should be a rapid drop after lag 1.

AR/MA (2) Model:

ACF: For an AR(2) process, you will see significant spikes at lags 1 and 2 in the ACF plot, and the autocorrelations for all other lags will be close to zero.

PACF: The PACF plot will have significant spikes at lags 1 and 2, and all other lags will be close to zero. There should be a rapid drop after lag 2.

AR/MA(3) Model:

ACF: In the ACF plot of an AR(3) process, there will be significant spikes at lags 1, 2, and 3, and the autocorrelations for all other lags will be close to zero.

PACF: The PACF plot will have significant spikes at lags 1, 2, and 3, and all other lags will be close to zero. There should be a rapid drop after lag 3.