SDA ASSIGNMENT - 2

Group No : 33

Team Details :

M Sai Amulya - S20170010099

G Sreeja - S20170010047

P Deepika Sowmya - S20170010110

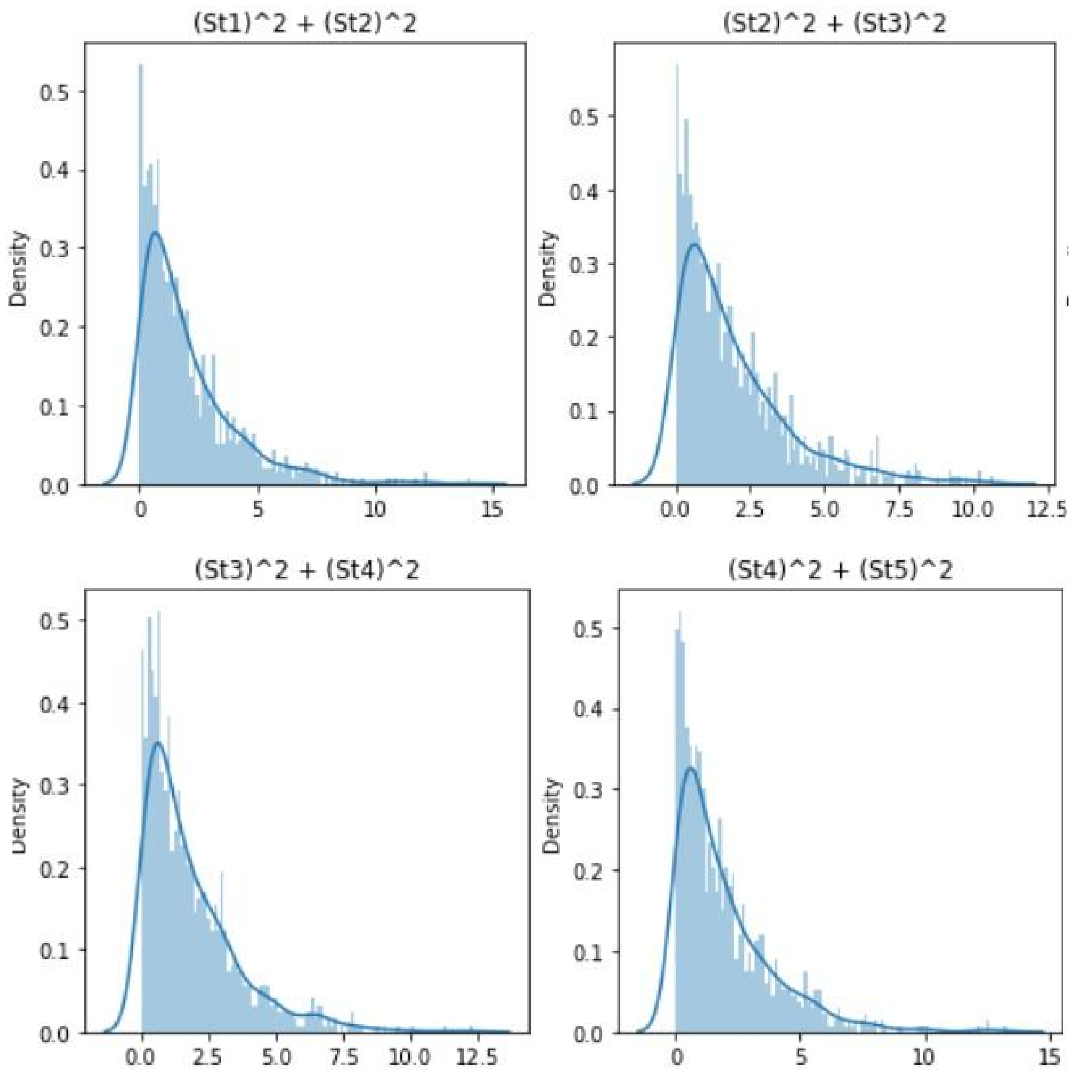
B Gayatri Shivani - S20170010029

1. Chi-Square Distribution

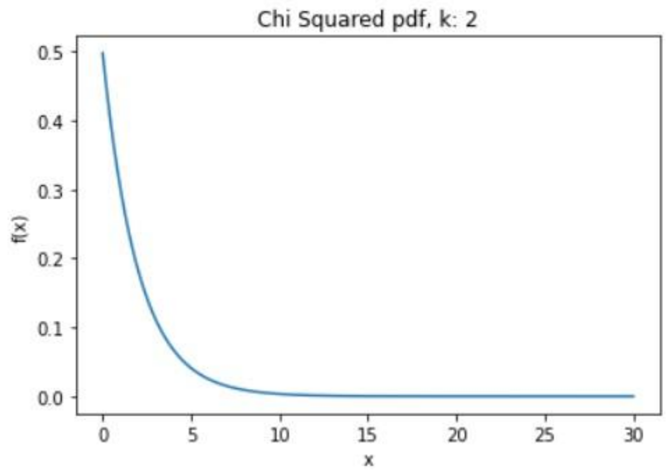
The chi-square distribution (χ^2-distribution) with k degrees of freedom is the distribution of a sum of the squares of k independent standard normal random variables.

Results :

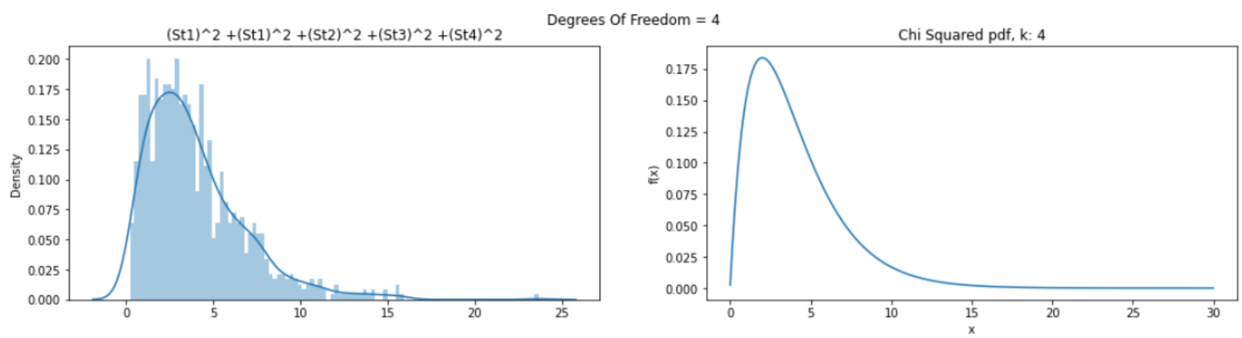
Different possibilities of $st(i)^2+st(j)^2$ are plotted below

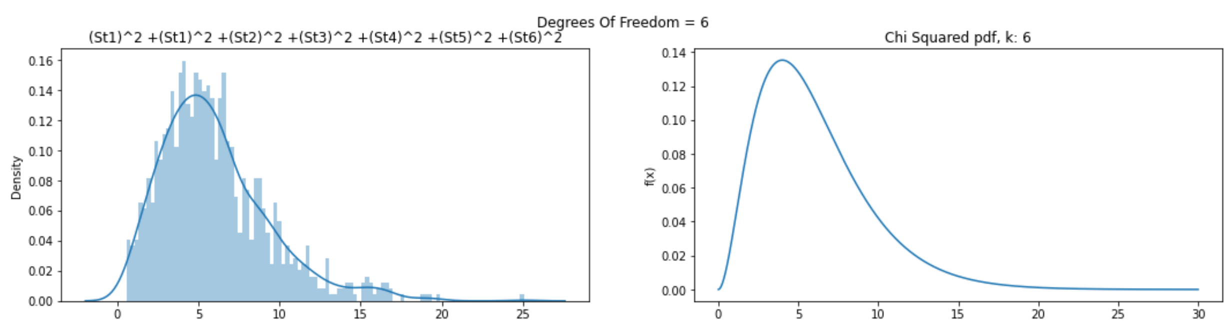


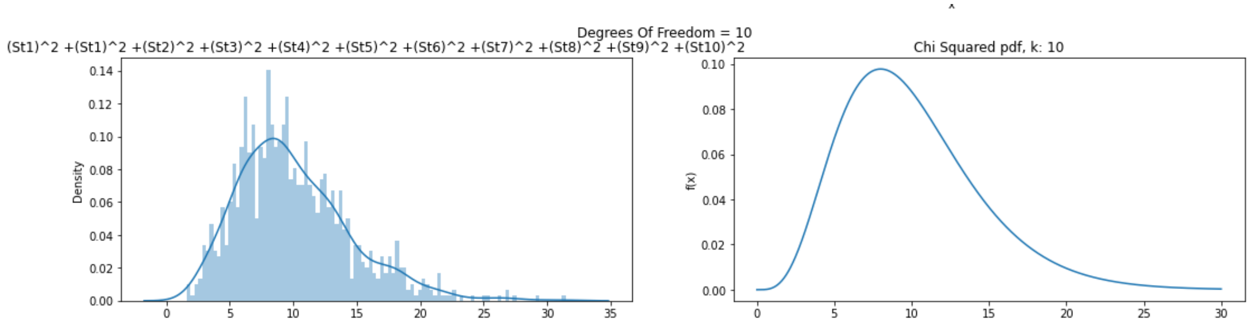
Chi squared pdf plot for degree of freedom : 2



Graphs are plotted for different degrees of freedom values by using population samples and chi square pdf







Conclusion :

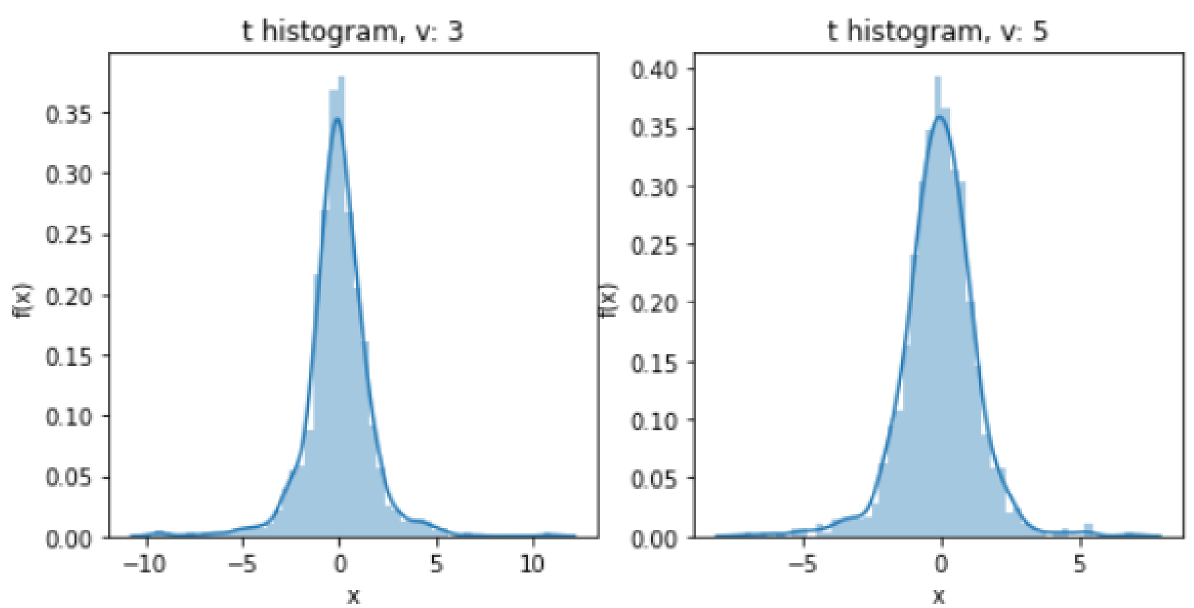
- Chi square distributions are always right skewed.
- The greater the degrees of freedom, the more the chi square distribution looks like a normal distribution.

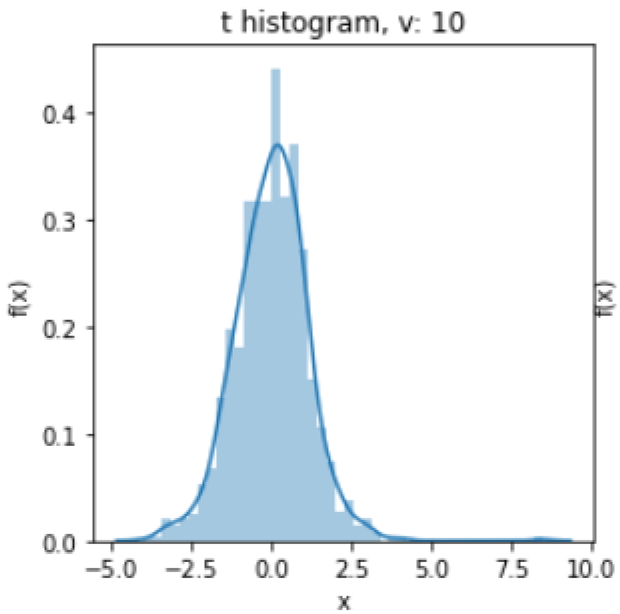
2. T - Distribution

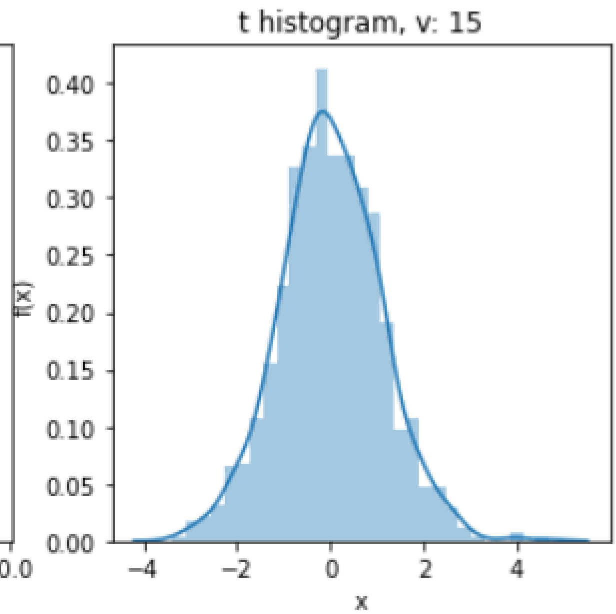
If we take a sample of n observations from a normal distribution, then the *t*-distribution with n-1 degrees of freedom can be defined as the distribution of the location of the sample mean relative to the true mean, divided by the sample standard deviation, after multiplying by the standardizing term √n

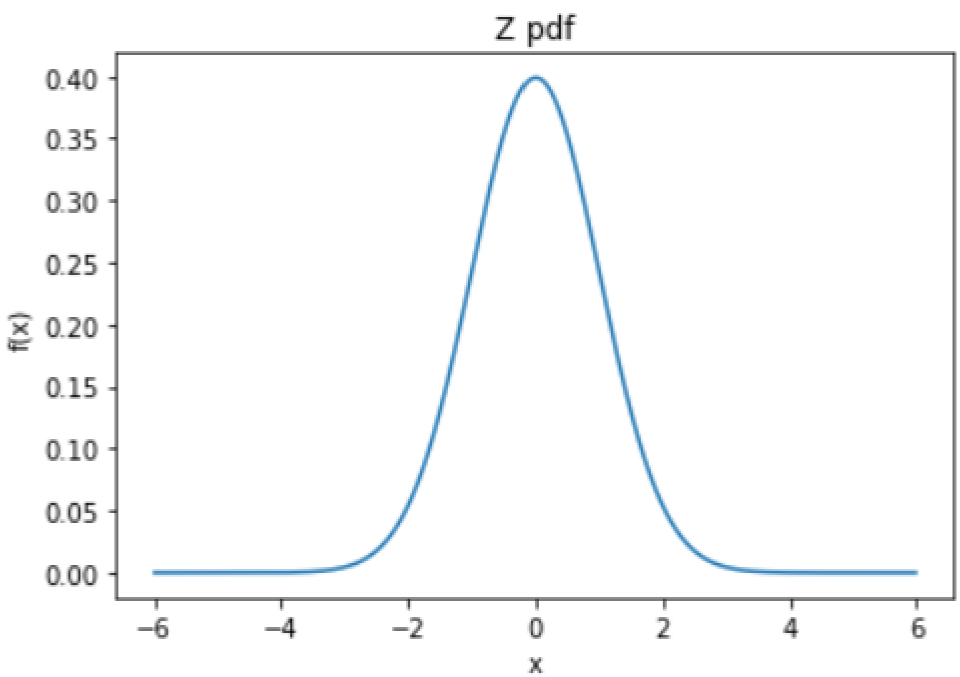
Results :

Graphs are plotted for different degrees of freedom values by using population samples and t-distribution pdf and the z graph is plotted as well.









Conclusion :

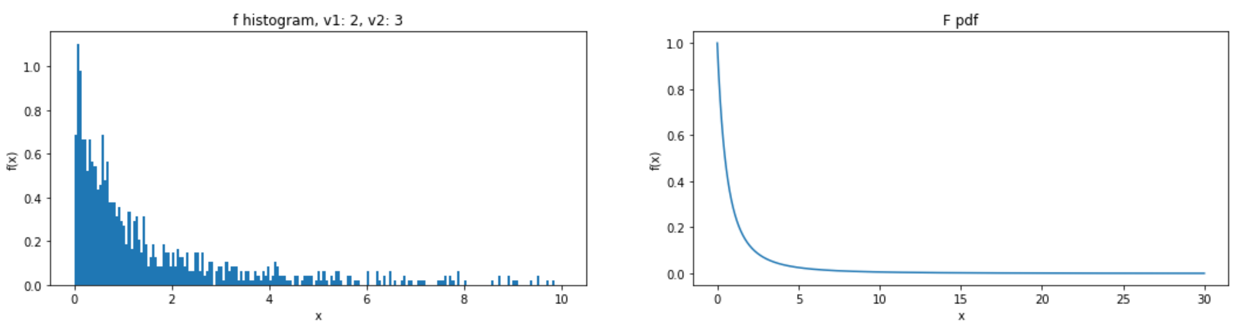
- T - distribution has greater chance at extremes than normal distribution due to heavier tails i.e., it tends to produce values far from mean.
- Small value of degree of freedom has heavier tails.
- WIth high values of degree of freedom it resembles normal distribution.

3. F - Distribution

The F distribution is the ratio of two chi-square distributions with degrees of freedom v_1 and v_2, respectively, where each chi-square has first been divided by its degrees of freedom.

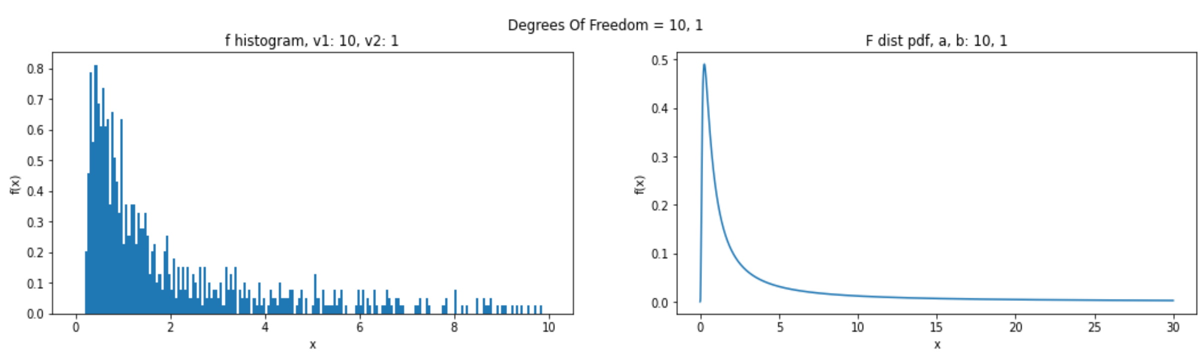
Results :

F-distribution pdf plot for degree of freedom : 2, 3

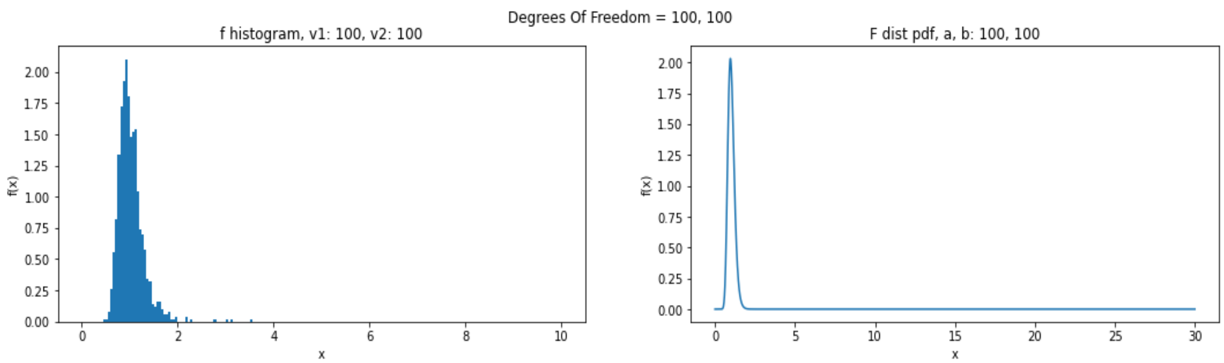


Graphs are plotted for different degrees of freedom values by using population samples and f-distribution pdf.

F-distribution pdf plot for degree of freedom : 10,1



F-distribution pdf plot for degree of freedom : 100,100



Conclusion :

- The curve is not symmetrical but skewed to the right.
- There is a different curve for each set of *df*s.
- As the degrees of freedom for the numerator and for the denominator get larger, the curve approximates the normal.

4. Multivariate Normal Data

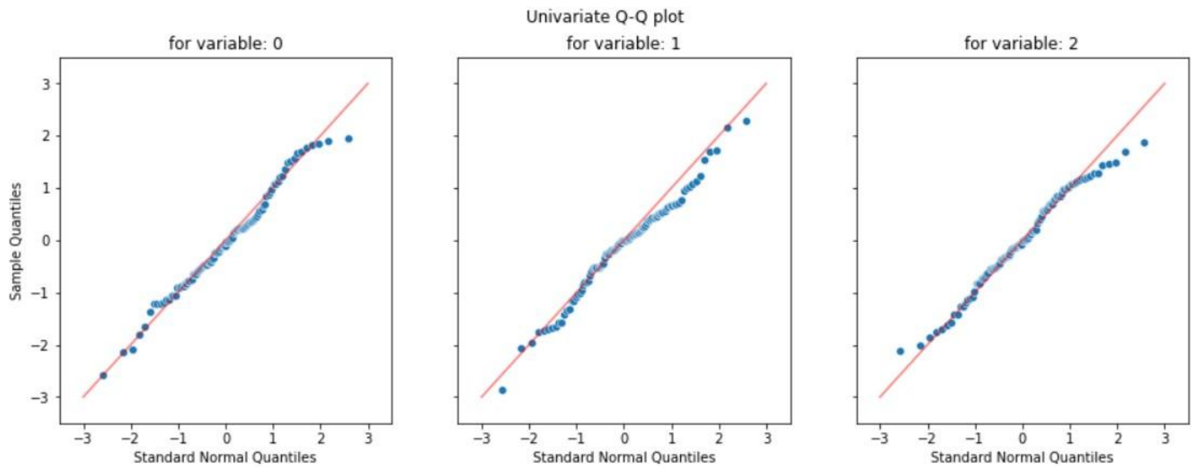
Multivariate normal data for 3 variables and 100 observations are taken with

- Mean = [0,0,0]
- Covariance = [[1,0,0],[0,1,0],[0,0,1]]

Univariate Marginal Test by Q-Q plot :

For each variable,we sorted the data and calculated standard normal quantiles for each observation with probability = (i-1/2)/(no of observations) for i^{th} quartile.

We have plotted the sorted observations with their standard normal quantile of the same probability.



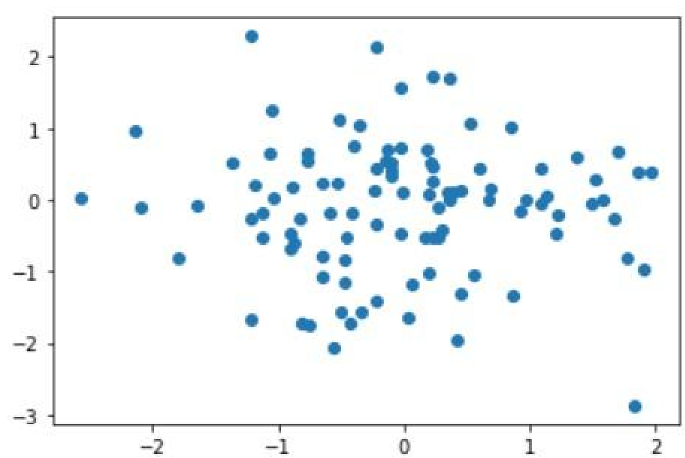
Conclusion :

Almost all the points lie on the same line. Hence all the 3 variables follow univariate normality.

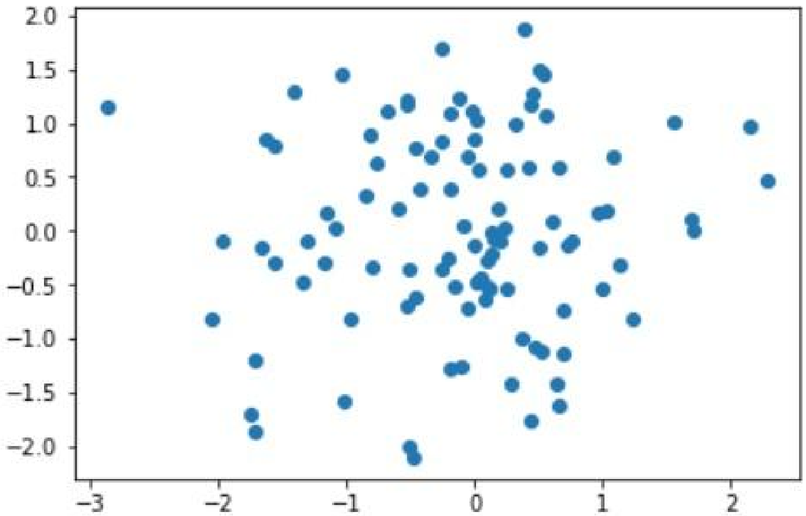
Bivariate and Multivariate Normality :

For bivariate normality, we plotted data for 2 variables of 100 observations in a graph to observe the distribution.

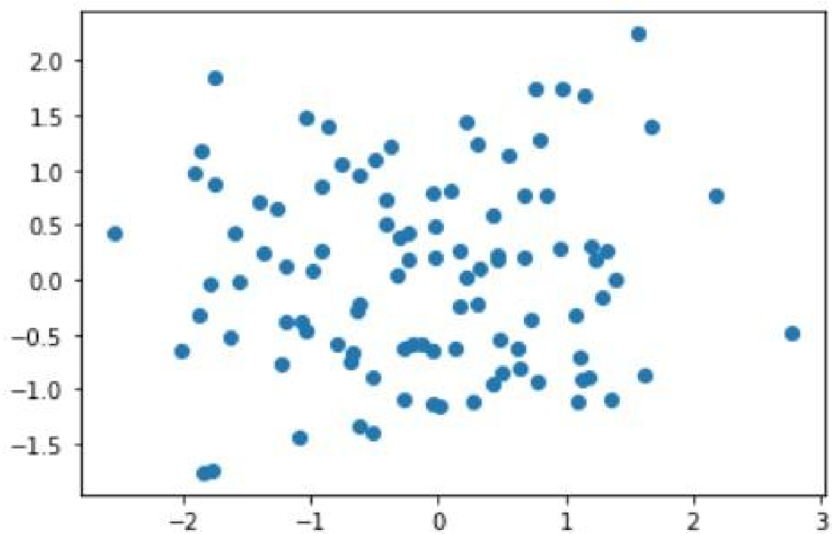
- Bivariate Normality of x and y



- Bivariate Normality of y and z



- Bivariate Normality of x and z



Conclusion:

As we observed all the plots formed are in elliptical shape which is a proof for bivariate normality.

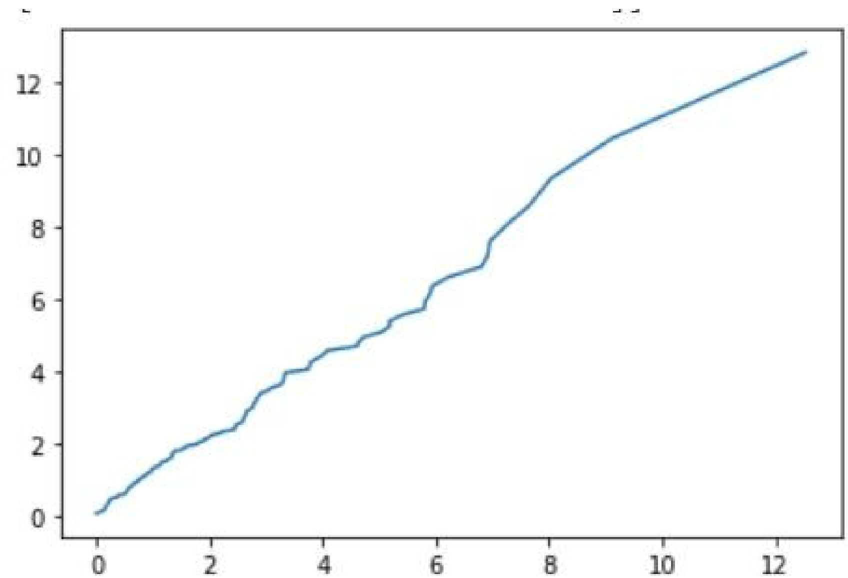
Multivariate Normality:

For Multivariate normality, We calculate dj square for every observation.

$d_j^2 = (x_j - \bar{x})' S^{-1} (x_j - \bar{x})$

here x_j is the jth observation $\bar{x}$ is mean and S is covariance.

We have sorted the data and calculated Chi-Squared quantile values with the probability (j-1/2)/(no of samples) for jth quantile.



Conclusion:

As we have observed most of the data points lie on the same line. Hence Multivariate Normality is proved.