SDA ASSIGNMENT-2

Group-44

Lushaank K - 20170010082

Vineesh S - 20170010132

Hemanth P G - 20170010048

Sumanth J - 20170010135

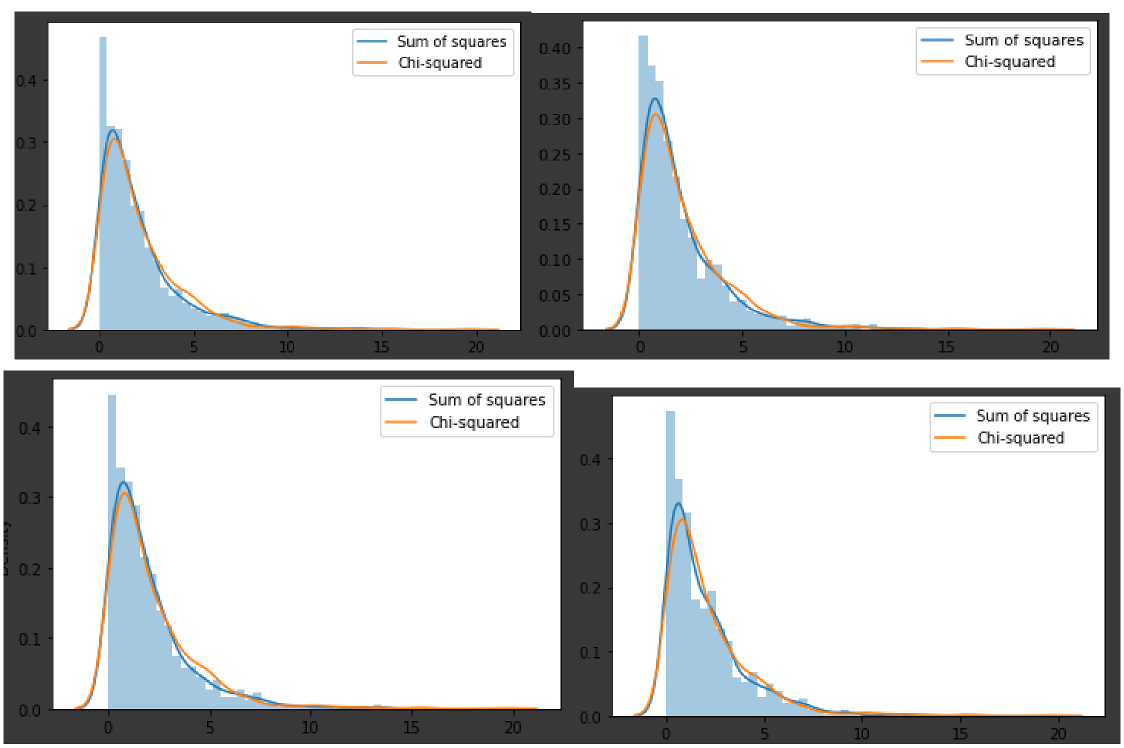
Question 1

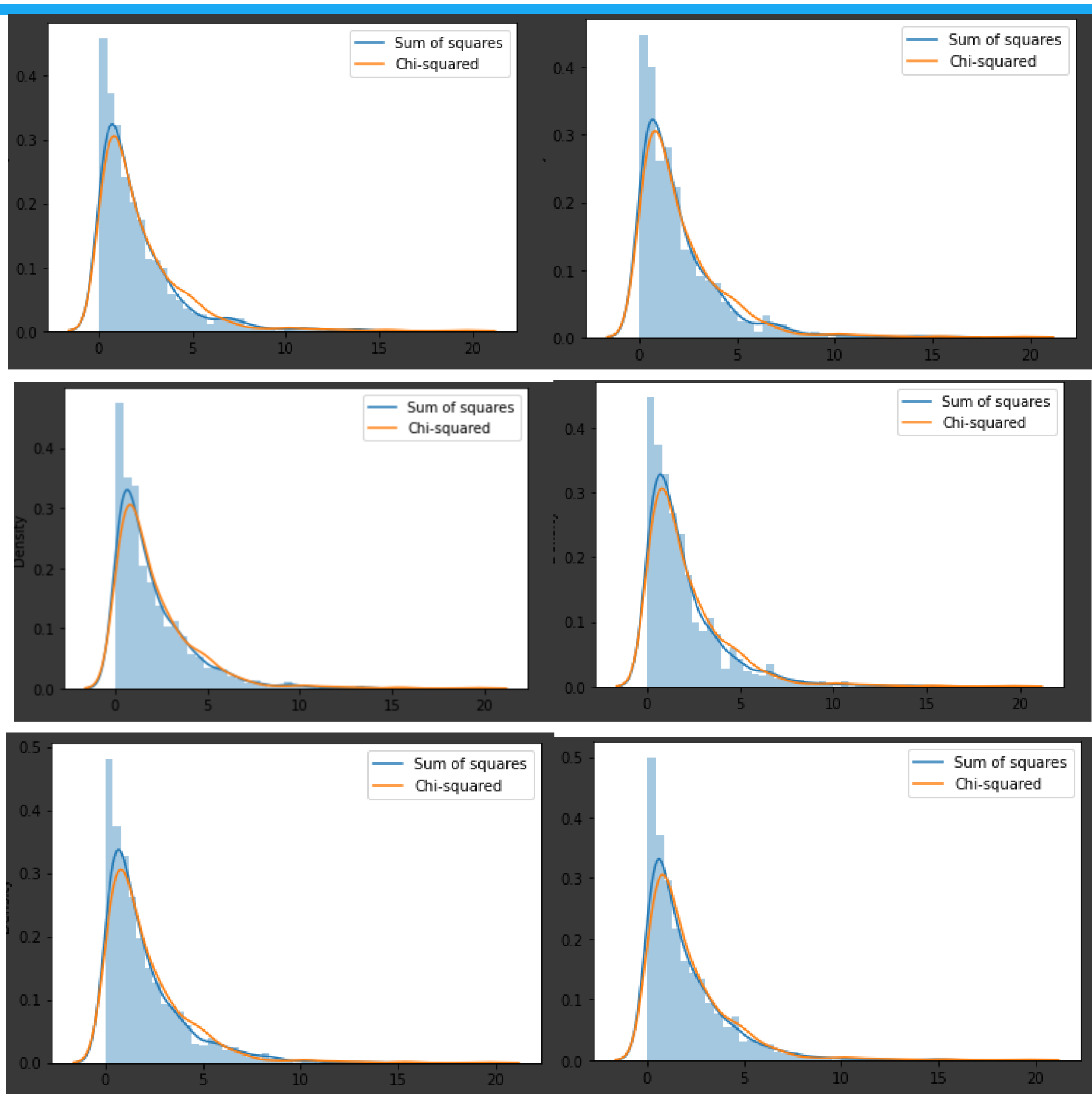
Taking random samples from standard normal distribution and finding the sum of the squares of the samples

Each sample consists of 1000 samples

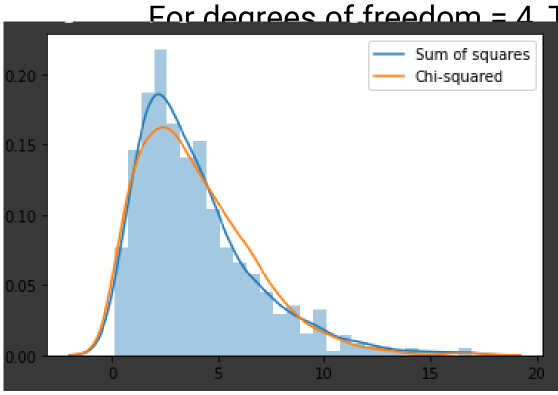
No of samples = 5, For degrees of freedom = 2

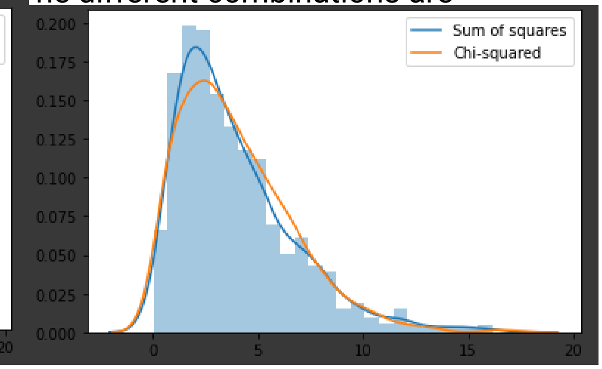
The different combinations are

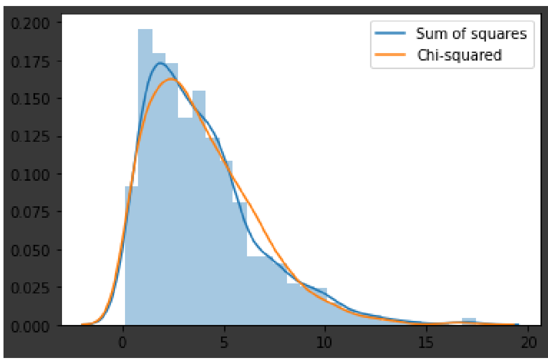


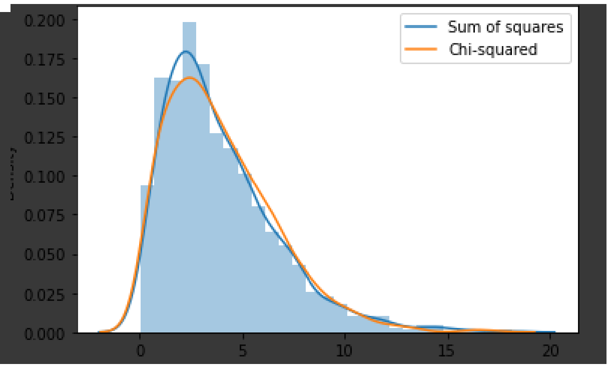

Sum of squares
Chi-squared
Sum of squares
Chi-squared
Density
Sum of squares
Chi-squared
Sum of squares
Chi-squared
Sum of squares
Chi-squared
Sum of squares
Chi-squared

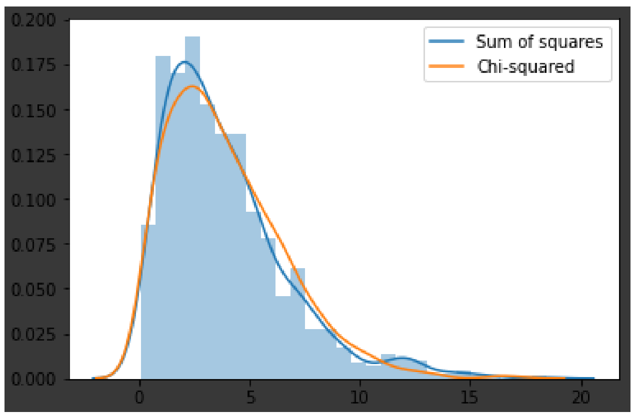
For degrees of freedom = 4, The different combinations are









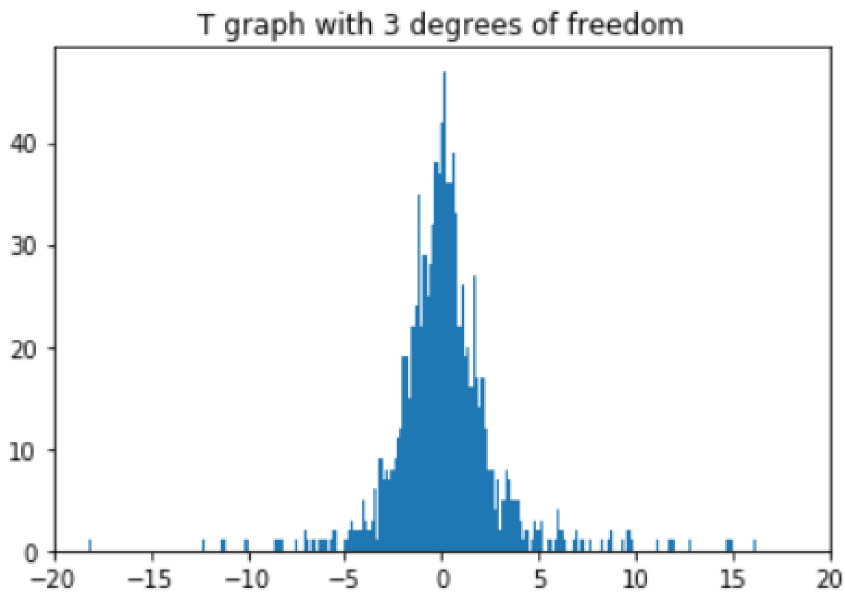


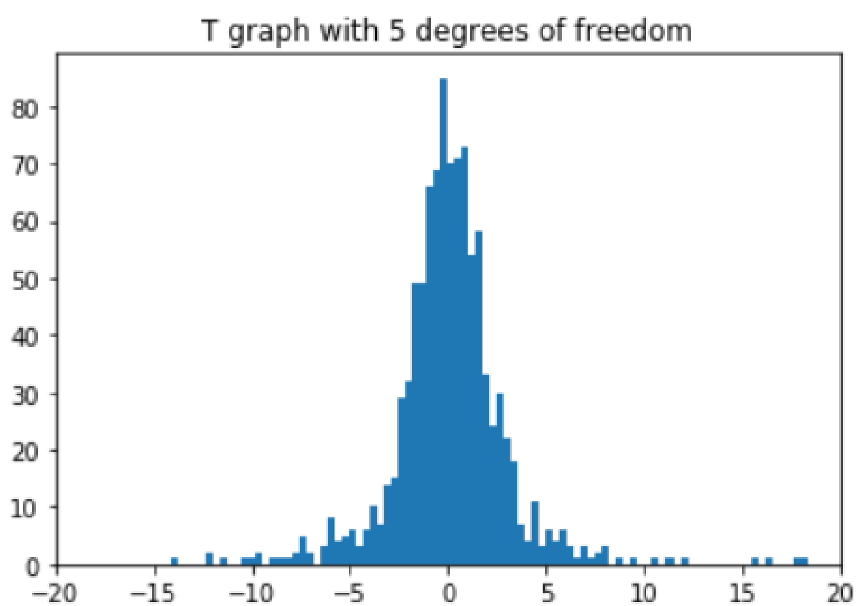
From the above plots we can conclude that sum of the squares of samples from normal distribution is approximately equal to the chi-squared distribution with same degrees of freedom

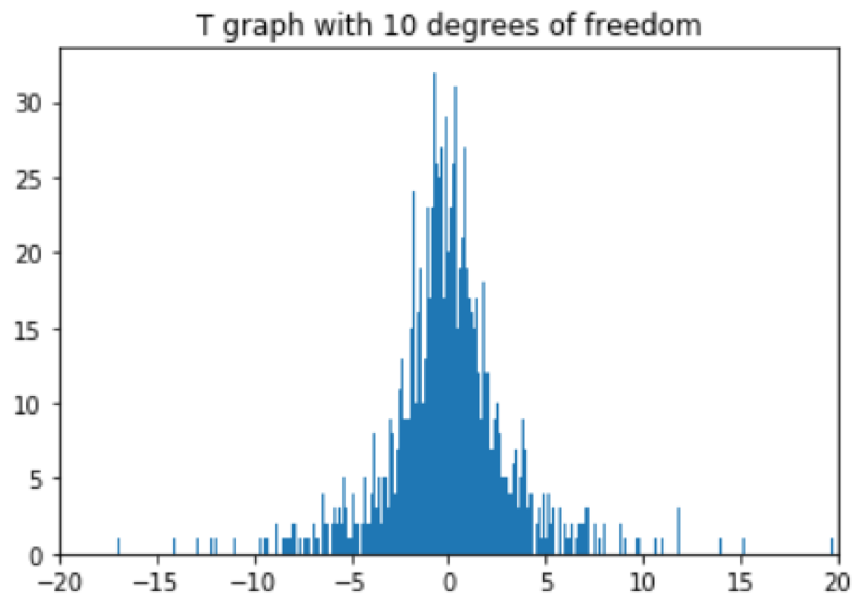
Question 2

Taking a random variable "t" such that t = Z/sqrt(V/v) in which Z is Normal Distribution with mean 0 and variance of 1 and V is Chi-Squared Distribution with 'v' degrees of freedom.

We have plotted different histograms for different degrees of freedom, namely







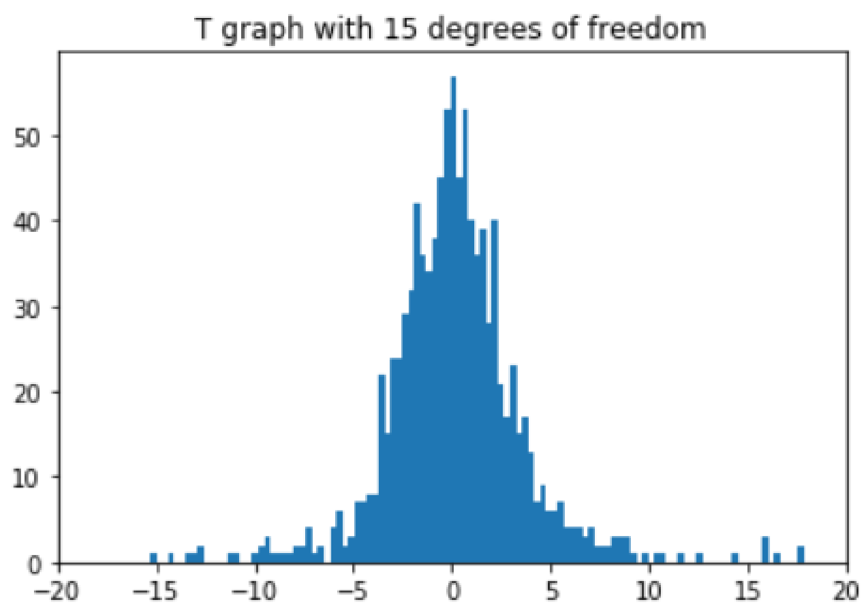


Fig:2.1-2.4: Shows the t graphs at different degrees of freedom namely 3, 5, 10, 15 respectively

Here's the Z graph with the same number of observations taken with comparison to t-graphs.

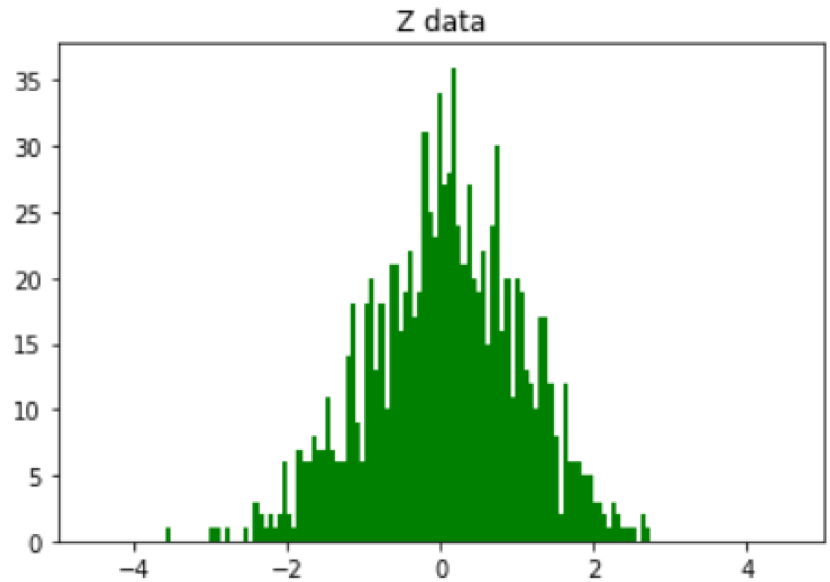


Fig:2.5: Shows the original Z graph taken 1000 samples

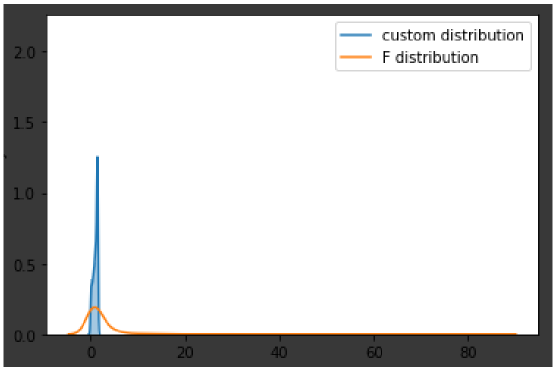
This shows us that for different degrees of freedom t - value actually taller at 0 for lower degrees of freedom and at higher degrees of freedom, the peak is slightly lower and could be divided more in the range.

All of these t graphs could be normally distributed and could be seen with comparison to the data we got from Normal Distribution. The data could be seen as more distributed when seen compared with this graph.

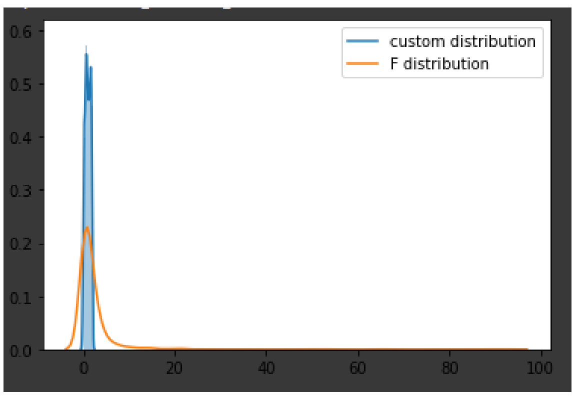
Question 3

F(n1,n2)= (U1/n1)/(U2/n2); U1 has χ2 distribution with n1 degrees of freedom. U2 has χ2 distribution with n2 degrees of freedom.

The plot for above function with n1=2 and n2=3 is



The plot for n1=2, n2=4



We can observe that this function is apporimately equal to F distribution with n1,n2 as parameters

Question 4

We simulate a multivariate data containing three variables having 100 observations. We plot the Quantile-Quantile plot as a test for normality for the Multivariate Data with Normal data specifically the Univariate Marginal Test.

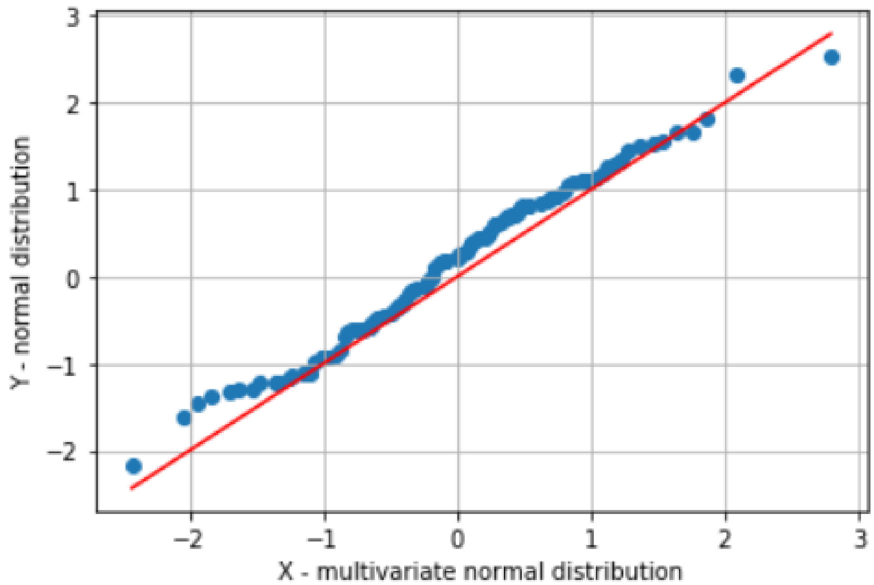


Fig 4.1: Q-Q Plot for checking normality

The straight plot shows a 45 degree along the x and y axis, that shows us the data we got is actually normal distribution. If we observe closely, we could say that the graph is slightly left skewed or we could say it is light tailed.

We could use the Shapiro-Wilks Test for checking normality which is generally one of the three general normality tests used to detect normality in Multivariate Normality. This test actually tests the null hypothesis that a sample x1, x2, x3, ... (here we use 100 samples for the test) from a normally distributed population. We use the test statistic

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2},$$

Fig 4.2: Test statistic

Here x(i), is i th order statistic and “x bar” is the mean of all the samples.

$$C = \|V^{-1}m\| = (m^{\top}V^{-1}V^{-1}m)^{1/2} \quad (a_1, \dots, a_n) = \frac{m^{\top}V^{-1}}{C}$$

Fig 4.3, 4.4: Shows how we compute coefficients

Then, when we use the null hypothesis and rejects model if $p < 0.05$, saying the model isn't normal.