**1.A F&B manager wants to determine whether there is any significant difference in the diameter of the cutlet between two units. A randomly selected sample of cutlets was collected from both units and measured? Analyze the data and draw inferences at 5% significance level. Please state the assumptions and tests that you carried out to check validity of the assumptions.**

**Minitab File : Cutlets.mtw**

Solution:

**#Ho: There is no difference in the diameter of the cutlets**

**#Ha: There is a difference in the diameter of the cutlets**

At 5% level of significance or 95% confidence level

The sample size of cutlet data set is of the size 34 with columns Unit\_A and Unit\_B

To check the normality of the Unit A and Unit B, we calculate the p-value

The p-value of Unit A is 0.31 which is greater than 0.05.

The p-value of Unit B is 0.522 which is greater than 0.05

**Both the units are normally distributed**

**Calculating the variance of the dataset:**

The p-value is 0.41. This means that the dataset has equal variances

**Calculating Two Sample T-Test for unequal Variance:**

scipy.stats.ttest\_ind(Cutlets.Unit\_A,Cutlets.Unit\_B,equal\_var = True)

Out[74]: Ttest\_indResult(statistic=0.7228688704678061, pvalue=0.4722394724599501)

From the above result the p-value is 0.47. This means that there is not enough sample evidence to reject the null hypothesis i.e Ho is not rejected.

**Conclusion:** There is no difference in the diameter of the cutlets

**2. A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. TAT is defined as sample collected to report dispatch.**

**Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level.**

Solution:

**#Ho: The means of Lab1, Lab2, Lab3 and Lab4 are equal**

**#Ha: The means of the different labs are not equal**

At 5% level of significance or 95% confidence level

The sample size of LabT data set is of the size 120 with columns Laboratory1, Laboratory2, Laboratory3 and Laboratory4

To check the normality of the Laboratory1, Laboratory2, Laboratory3 and Laboratory4, we calculate the p-value

The p-value of Laboratory1 is 0.5506953597068787 which is greater than 0.05.

The p-value of Laboratory2 is 0.8637524843215942 which is greater than 0.05

The p-value of Laboratory3 is 0.4205053448677063 which is greater than 0.05

The p-value of Laboratory2 is 0.6618951559066772 which is greater than 0.05

From the above p-values there is no evidence to reject that the data of all the above labs are not normally distributed. Hence, the data of all features/ variables are normally distributed.

**Calculating the variance of the datasets between Labs:**

There are four labs which means that there are 6 combinations of 2 to test the variance between Laboratories.

scipy.stats.levene(cof.Laboratory1, cof.Laboratory2)

Out[31]: LeveneResult(statistic=3.5495027780905763, pvalue=0.06078228171776711)

There is **no** variance between Lab1 and Lab2

scipy.stats.levene(cof.Laboratory1, cof.Laboratory3)

Out[32]: LeveneResult(statistic=7.547664894290509, pvalue=0.006468575869839467)

There is a variance between Lab1 and Lab2

scipy.stats.levene(cof.Laboratory1, cof.Laboratory4)

Out[33]: LeveneResult(statistic=1.5000140718506723, pvalue=0.22188001348277267)

There is **no** variance between Lab1 and Lab4

scipy.stats.levene(cof.Laboratory2, cof.Laboratory3)

Out[34]: LeveneResult(statistic=0.9441465124387124, pvalue=0.33220021420602397)

There is **no** variance between Lab2 and Lab3

scipy.stats.levene(cof.Laboratory2, cof.Laboratory4)

Out[35]: LeveneResult(statistic=0.2889202799636133, pvalue=0.5914154837597723)

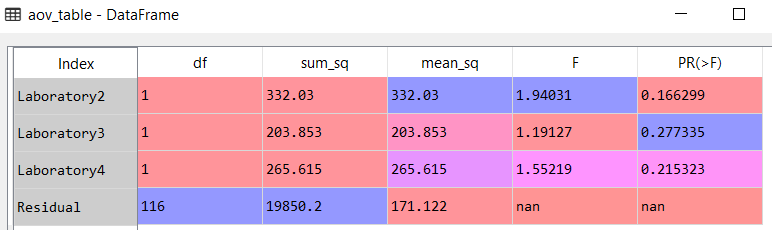
There is **no** variance between Lab2 and Lab4

scipy.stats.levene(cof.Laboratory3, cof.Laboratory4)

Out[36]: LeveneResult(statistic=2.037958464521512, pvalue=0.15472618294425391)

There is **no** variance between Lab3 and Lab4

The following are the results of one-way annova test between laboratories to evaluate if the means are equal or unequal:



From the above table the p-values are greater than 0.05 (considering 5% level of significance).

**There is no sample evidence to reject null hypothesis Ho**

Conclusion: This means that the average Turn Around Time(TAT) of reports of Labs have no differences.

**3. Sales of products in four different regions is tabulated for males and females. Find if male-female buyer rations are similar across regions.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **East** | **West** | **North** | **South** |
| **Males** | **50** | **142** | **131** | **70** |
| **Females** | **550** | **351** | **480** | **350** |

Solution:

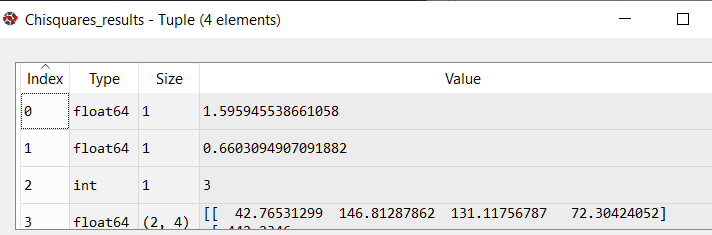
Ho= Proportions of Male and Female are same

Ha= Proportions of Male and Female are not same

At 5% level of significance or 95% confidence level

The dimensions of the matrix is (2,4)

Executing Chi-square test using scipy.stats, we get the following results:



The p-value here is 0.66030 which is greater than 0.05. This means that there is no enough sample evidence to reject the null hypothesis.

**Conclusion:** The proportions of Male and Female are same