1: **Hypothesis Testing Exercise**

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**

ANS : import pandas as pd

import numpy as np

from scipy import stats

from scipy.stats import norm

data=pd.read\_csv('D://Cutlets.csv')

data

unitA=pd.Series(data.iloc[:,0])

unitA

unitB=pd.Series(data.iloc[:,1])

unitB

p\_value=stats.ttest\_ind(unitA,unitB)

p\_value

p\_value[1]

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.

Minitab File: **LabTAT.mtw**

ANS: import pandas as pd

import numpy as np

from scipy import stats

from scipy.stats import norm

data=pd.read\_csv('D:\\LabTAT.csv')

data.head()

Laboratory 1 Laboratory 2 Laboratory 3 Laboratory 4

0 185.35 165.53 176.70 166.13

1 170.49 185.91 198.45 160.79

2 192.77 194.92 201.23 185.18

3 177.33 183.00 199.61 176.42

4 193.41 169.57 204.63 152.60

p\_value=stats.f\_oneway(data.iloc[:,0],data.iloc[:,1],data.iloc[:,2],data.iloc[:,3])

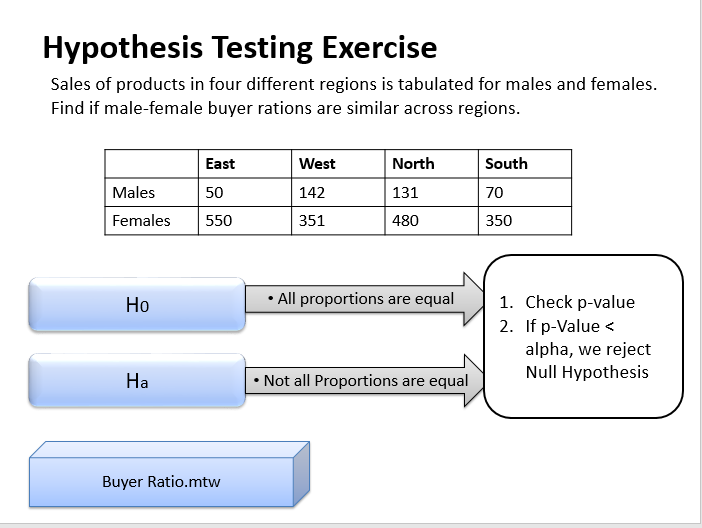
p\_value

F\_onewayResult(statistic=118.70421654401437, pvalue=2.1156708949992414e-57)

p\_value[1]

2.1156708949992414e-57

​

3: 

ANS : import numpy as np

import scipy.stats as stats

import pandas as pd

df=pd.read\_csv('E:\\assiment\\BuyerRatio.csv')

df

Observed Values East West North South

0 Males 50 142 131 70

1 Females 435 1523 1356 750

table = [[50, 142, 131, 70],

[435, 1523, 1356, 750]]

print(table)

[[50, 142, 131, 70], [435, 1523, 1356, 750]]

stat, p, dof,expected = stats.chi2\_contingency(table)

print('dof=%d' % dof)

print(p)

dof=3

p=0.6603094907091882

alpha = 0.05

print('significance=%.3f, p=%.3f' % (alpha, p))

if p <= alpha:

print('Dependent (reject H0)')

else:

print('Independent (fail to reject H0)')

significance=0.050, p=0.660

Independent (fail to reject H0)

4: TeleCall uses 4 centers around the globe to process customer order forms. They audit a certain % of the customer order forms. Any error in order form renders it defective and has to be reworked before processing. The manager wants to check whether the defective % varies by centre. Please analyze the data at *5%* significance level and help the manager draw appropriate inferences

Minitab File: **CustomerOrderForm.mtw**

 ANS : import pandas as pd

import numpy as np

from scipy import stats

from scipy.stats import norm

from scipy.stats import chi2\_contingency

data=pd.read\_csv('E:\\assiment\\Costomer+OrderForm.csv')

data.head()

data.Phillippines.value\_counts()

data.Indonesia.value\_counts()

data.Malta.value\_counts()

data.India.value\_counts()

obs=np.array([[271,267,269,280],[29,33,31,20]])

obs

chi2\_contingency(obs)