Importing Libraries

In [2]:

```python
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [3]:

```python
from scipy.stats import chi2 # Distribution (cdf etc.)
from scipy.stats import chisquare # Statistical test (chistat, pvalue)
from scipy.stats import chi2_contingency # Categorical Vs Categorical
from scipy.stats import ttest_rel,ttest_1samp,ttest_ind
from scipy.stats import binom,tiecorrect
```

Chi Square Tests

In [4]:

```python
# Ho : Coin is fair
# Ha : Coin is Biased

chi_stat=((28-25)**2/25)+((22-25)**2/25)
print("chi_stat : ",chi_stat)
p_value = 1-chi2.cdf(chi_stat,df=1)
print("p_value : ",p_value)
alpha = 0.05

if p_value<alpha:
    print("Interpretation : Reject Ho")
else:
    print("Interpretation : Fail to Reject Ho")
```

```
chi_stat :  0.72
p_value :  0.3961439091520741
Interpretation : Fail to Reject Ho
```

In [6]:

```python
# Ho : Coin is fair
# Ha : Coin is Biased

chi_stat,p_value=chisquare([28,22],[25,25])
print("chi_stat : ",chi_stat)
print("p_value : ",p_value)
alpha = 0.05

if p_value<alpha:
    print("Interpretation : Reject Ho")
else:
    print("Interpretation : Fail to Reject Ho")
```

```
chi_stat :  0.72
p_value :  0.3961439091520741
Interpretation : Fail to Reject Ho
```

In [7]:

```python
# Ho : Coin is fair
# Ha : Coin is Biased

chi_stat,p_value=chisquare([45,5],[25,25])
print("chi_stat : ",chi_stat)
print("p_value : ",p_value)
alpha = 0.05

if p_value<alpha:
    print("Interpretation : Reject Ho")
else:
    print("Interpretation : Fail to Reject Ho")
```

```
chi_stat :  32.0
p_value :  1.5417257900280013e-08
Interpretation : Reject Ho
```

In [8]:

```python
critical_chi_stat= chi2.ppf(0.95,df=1)
print("critical_chi_stat : ",critical_chi_stat)
```

```
critical_chi_stat :  3.841458820694124
```

In [ ]:

In [13]:

```python
# Ho : Bulb manufacturer is right
# Ha : Bulb manufacturer is wrong

chi_stat,p_value=chisquare([60,10,10,30,50,20,20],[50,20,30,10,40,30,20])
print("chi_stat : ",chi_stat)
print("p_value : ",p_value)
alpha = 0.05

if p_value<alpha:
    print("Interpretation : Reject Ho")
else:
    print("Interpretation : Fail to Reject Ho")
```

```
chi_stat :  66.16666666666667
p_value :  2.4918317217002336e-12
Interpretation : Reject Ho
```

In [14]:

```python
chisquare?
```

Test of Independance

In [ ]:

In [11]:

```python
# Ho : Gender doesnt affect the buying pattern ( Independant)
# Ha : Gender affects the buying pattern (dependant)

observed = [[527,72],[206,102]]
chi_stat,p_value,dof,expected_freq=chi2_contingency(observed)
print("chi_stat : ",chi_stat)
print("p_value : ",p_value)
print("dof : ",dof)
print("expected_freq : ",expected_freq)
alpha = 0.05

if p_value<alpha:
    print("Interpretation : Reject Ho")
else:
    print("Interpretation : Fail to Reject Ho")
```

```
chi_stat :  57.04098674049609
p_value :  4.268230756875865e-14
dof :  1
expected_freq :  [[484.08710033 114.91289967]
 [248.91289967  59.08710033]]
Interpretation : Reject Ho
```

In [15]:

```
!ls
```

```
06_Chi_Squared_Test_Notebook.ipynb iq_two_schools.csv
Sachin_ODI.csv                     problem_solving.csv
aerofit.csv
```

In [16]:

```python
df=pd.read_csv("aerofit.csv")
```

In [17]:

```python
df
```

Out[17]:

	Product	Age	Gender	Education	MaritalStatus	Usage	Fitness	Income	Miles
0	KP281	18	Male	14	Single	3	4	29562	112
1	KP281	19	Male	15	Single	2	3	31836	75
2	KP281	19	Female	14	Partnered	4	3	30699	66
3	KP281	19	Male	12	Single	3	3	32973	85
4	KP281	20	Male	13	Partnered	4	2	35247	47
...	...	...	...	...	...	...	...	...	...
175	KP781	40	Male	21	Single	6	5	83416	200
176	KP781	42	Male	18	Single	5	4	89641	200
177	KP781	45	Male	16	Single	5	5	90886	160
178	KP781	47	Male	18	Partnered	4	5	104581	120
179	KP781	48	Male	18	Partnered	4	5	95508	180

180 rows × 9 columns

In [20]:

```python
gender_product=pd.crosstab(index=df["Gender"],columns=df["Product"])
gender_product
```

Out[20]:

Product	KP281	KP481	KP781
Gender			
Female	40	29	7
Male	40	31	33

In [21]:

```python
# Ho : Gender doesnt affect the product choice( Independant)
# Ha : Gender affects the product choice (dependant)

observed = gender_product
chi_stat,p_value,dof,expected_freq=chi2_contingency(observed)
print("chi_stat : ",chi_stat)
print("p_value : ",p_value)
print("dof : ",dof)
print("expected_freq : ",expected_freq)
alpha = 0.05

if p_value<alpha:
    print("Interpretation : Reject Ho")
else:
    print("Interpretation : Fail to Reject Ho")
```

```
chi_stat :  12.923836032388664
p_value :  0.0015617972833158714
dof :  2
expected_freq :  [[33.77777778 25.33333333 16.88888889]
 [46.22222222 34.66666667 23.11111111]]
Interpretation : Reject Ho
```

In [ ]:

In [ ]:

```python
# Product vs Income (pd.cut -->  bins --> chi2_contingency)
```

In [ ]:

```python
# Gender vs Income (pd.cut -->  bins --> chi2_contingency)
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

```
In [ ]:
```

```
In [ ]:
```