Practical No.4

Aim:- Hypothesis Testing

- Formulate null and alternative hypotheses for a given problem.
- Conduct a hypothesis test using appropriate statistical tests (e.g., t-test, chi-square test).
- Interpret the results and draw conclusions based on the test outcomes.

In [1]:
```
from scipy import stats
import numpy as np
```

One Sampled T-Test

In [2]:
```
#Creating a sample of ages
ages = [45,89,23,46,12,69,45,24,34,67]
print(ages)

#Calculating the mean of the sample
mean = np.mean(ages)
print("Mean of Age values : ",mean)

#Define the null hypothesis
H0 = "The average age of 10 people is 30."

#Define the alternative hypothesis
H1 = "The average age of 10 people is  more than 30."

#Perforimg the T-Test
t_stats,p_val = stats.ttest_1samp(ages,30)
print("P-value is: ",p_val)
print("The T-Statistics is: ",t_stats)

#taking the threshold value as 0.05 or 5%
if p_val < 0.05:
    print("We can reject the null hypothesis")
else:
    print("We can accept the null hypotheis")
```

```
[45, 89, 23, 46, 12, 69, 45, 24, 34, 67]
Mean of Age values :  45.4
P-value is:  0.07179988272763561
The T-Statistics is:  2.0397003109502543
We can accept the null hypotheis
```

Independent T-Test or Two Sampled T-Test

In [3]:
```
#Creating the data groups
data_group1 = np.array([12, 18,12, 13,15,1,7,
                        20,21,25,19,31,21,17,
                        17,15,19,15,12,15])
data_group2 = np.array([23,22,24,25,21,26,21,
                          21,25,30,24,21,23,19,
                          14,18,14,12,19,15])
#Calculating the mean of the two data groups
mean1=np.mean(data_group1)
mean2=np.mean(data_group2)

#Print mean values
print("data group 1 mean value:",mean1)
print("data group 2 mean value:",mean2)

#Calculating standard deviation
std1=np.std(data_group1)
std2=np.std(data_group2)

#Printing standard deviation values
print("data group 1 std value:",std1)
print("data group 2 std value:",std2)
```

```
data group 1 mean value: 16.25
data group 2 mean value: 20.85
data group 1 std value: 6.171507109288622
data group 2 std value: 4.452808102759426
```

In [4]:
```
#Define the null hypothesis
H0 = "Independent sample means are equal."

#Define the alternative hypothesis
H1 = "Independent sample means are not equal."

#Implementing the Two-Sampled t-test or Independent Sampled t-test.
t_stats,p_val=stats.ttest_ind(data_group1,data_group2)
print("the P-val is: ",p_val)
print("the T-statistics is: ",t_stats)

#taking the threshold value as 0.05 or 5%
if p_val < 0.05:
    print("We can reject the null hypothesis")
else:
    print("We can accept the null hypotheis")
```

```
the P-val is:  0.012117171124028792
the T-statistics is:  -2.6347481110277466
We can reject the null hypothesis
```

In [5]:
```
#Creating Two samples
sample1=[29,30,33,41,38,36,35,31,29,30]
sample2=[31,32,33,39,30,33,30,28,29,31]

#Define the null hypothesis
H0 = "Dependent sample means are equal."

#Define the alternative hypothesis
H1 = "Dependent sample means are not equal."

#Performing paired sample t-test
t_stats,p_val=stats.ttest_rel(sample1,sample2)
print("the P-val is: ",p_val)
print("the T-statistics is: ",t_stats)

#taking the threshold value as 0.05 or 5%
if p_val < 0.05:
    print("We can reject the null hypothesis")
else:
    print("We can accept the null hypotheis")
```

```
the P-val is:  0.15266056244408904
the T-statistics is:  1.5622669317698863
We can accept the null hypotheis
```

Chi-Square Test

In [6]:
```
#defining data
data = [[231,256,321],[245,312,213]]

#Define the null hypothesis
H0 = "There is no relation between variables."

#Define the alternative hypothesis
H1 = "There is significiant relation between variables."

#Performing chi-square test
t_stats,p_val, dof,expected_val=stats.chi2_contingency(data)
print("The p-val of our test is: "+str(p_val))

#Checking the hypothesis
if p_val <= 0.05:
    print("We can reject the null hypothesis")
else:
    print("We can accept the null hypotheis")
```

```
The p-val of our test is: 1.4585823594475804e-06
We can reject the null hypothesis
```