Practical 8 : Hypothesis testing

Name : $Deepak\ Kumar\ Bharti$

Roll No. : 25056765013

Group : A

Perform hypothesis testing in the following cases using both **parametric** and **non-parametric** approaches.

Case 1 :

A coaching institute claims that the average score of students is 70 marks.

A random sample of 10 students obtained the following marks:

68, 72, 69, 71, 70, 67, 73, 74, 66, 69

Test the claim at 5% level of significance. (one-sample t-test & Wilcoxon Signed-Rank test.)

-: Methodology :-

For one sample t-test : (Parametric)

A **one-sample t-test** is used when we want to check whether the mean of a single sample is significantly different from a known or hypothesized population mean.

Hypotheses :

- **Null Hypothesis (H_0) :** μ = 70 (**i.e.,** The average score of students is 70 marks.)
- **Alternative Hypothesis (H_1) :** μ ≠ 70 (**i.e.,** The average score of students is not 70 marks.)

Test Statistic :

The *t* statistic is :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Where

- $\bar{x}$ = sample mean
- μ_0 = hypothesized mean
- s = sample standard deviation

- n = sample size

Degrees of freedom :

$$df = n - 1$$

Decision Rule :

If

$$|t| > t_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$$t_{critical} = t_{\frac{\alpha}{2},(n-1)} \text{ (i.e.,the tabulated value of t statistics)}$$

In [1]:
```
import numpy as np
from scipy import stats
from scipy.stats import t
from scipy.stats import wilcoxon
from scipy.stats import chi2
```

In [2]:
```
# Sample data
students_marks = np.array([68, 72, 69, 71, 70, 67, 73, 74, 66, 69])
mu = 70

# One-sample t-test
t1_stat, t1_pvalue = stats.ttest_1samp(students_marks, mu)

alpha_1 = 0.05
n1 = len(students_marks)
df_1 = n1 - 1

# two-tailed critical value
t1_critical = t.ppf(1 - alpha_1/2, df_1)

print("One-sample t-test")
print("="*20)
print(f"Degree of freedom : {df_1}")
print(f"Critical(Tabulated) t value : {t1_critical:.4f}")
print(f"Calculated t-statistic : {t1_stat:.4f}")
print(f"p-value : {t1_pvalue:.4f}")
print("\nDision :")
if t1_pvalue < 0.05 or abs(t1_stat) > t1_critical :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
One-sample t-test
====================
Degree of freedom : 9
Critical(Tabulated) t value : 2.2622
Calculated t-statistic : -0.1216
p-value : 0.9059

Dision :
Failed to reject the Null Hypothesis
```

Conclusion :

Here,

$$|t| = 0.1216 < 2.2622 = t_{critical}$$

or

$$p - value = 0.9095 > 0.05$$

So, we may **accept** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, the coaching institute claims that the average score of the students is 70 marks is **valid**.

-: Methodology :-

Wilcoxon signed-rank test : (Non-Parametric)

The **Wilcoxon signed-rank test** is a non-parametric alternative to the one-sample t-test (and paired t-test). We use it when we cannot assume normality but still want to test a median difference.

Step-by-step procedure

Step 1: State the hypothesis value

- **Null Hypothesis (H_0) :** μ = 70 (**i.e.,** The average score of students is 70 marks.)
- **Alternative Hypothesis (H_1) :** μ ≠ 70 (**i.e.,** The average score of students is not 70 marks.)

Step 2: Compute differences For each observation x_i :

$$d_i = x_i - \mu_0$$

Step 3: Remove zero differences

- Any $d_i = 0$ is dropped.
- Reduce sample size accordingly

Step 4: Take the absolute value of differences

$$|d_i|$$

Step 5: Rank the absolute differences

- Rank $|d_i|$ from smallest to largest.
- If ties occur then assign average ranks

Step 6: Assign signs to ranks (Give each rank the sign of its original difference)

- If $d_i > 0$ then rank is positive
- If $d_i < 0$ then rank is negative

Step 7: Compute the sum of ranks

- W^+ = sum of positive signed ranks
- W^- = sum of negative signed ranks

Step 8: Compute test Wilcoxon test statistic (W)

$$W = min(W^+, W^-) \text{ (i.e., the Wilcoxon signed-rank statistic)}$$

Step 6: Decision Rule :

If

$$W > W_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$$W_{critical} = W_{\alpha,n} \text{ (i.e.,the tabulated value of wilcoxon statistics W, for sample size n)}$$

In [3]:
```
# Wilcoxon signed-rank test
w1_stat, w1_pvalue = stats.wilcoxon(students_marks-70)

print("Wilcoxon signed-rank test")
print("="*25)
print(f"Wilcoxon statistic : {w1_stat:.4f}")
print(f"p-value : {w1_pvalue:.4f}")
print("\nDision :")
if w1_pvalue < 0.05 :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
Wilcoxon signed-rank test
=========================
Wilcoxon statistic : 21.5000
p-value : 0.9961

Dision :
Failed to reject the Null Hypothesis
```

Conclusion :

Here,

$$p - value = 0.9961 > 0.05$$

So, we may **accept** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, the coaching institute claims that the average score of the students is 70 marks is **valid**.

Case 2 :

Marks of students taught by two different methods :

Group A : 75, 78, 74, 77, 76, 73, 79, 80

Group B : 72, 70, 68, 71, 69, 73, 67, 70

Test whether the mean marks differ significantly. (t-test for independence & Mann–Whitney U test)

-: Methodology :-

---

t-test for independence : (Parametric)

The **independent t-test** is used to check whether the means of two independent groups are significantly different.

Hypotheses :

- **Null Hypothesis (H_0) :** $\mu_A = \mu_B$ (**i.e.,** There is no significant difference in the mean marks of students taught by the two methods.)
- **Alternative Hypothesis (H_1) :** $\mu_A \neq \mu_B$ (**i.e.,** There is a significant difference in the mean marks of students taught by the two methods.)

Where

- μ_A = population mean of Group A
- μ_B = population mean of Group B

Test Statistic :

- If population variances are assumed equal :

The t statistic is :

$$t = \frac{\bar{x}_A - \bar{x}_B}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$$

Where

- Where the pooled variance (s_p^2) is:

$$s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}$$

and

- $\bar{x}_A$ = sample mean of Group A
- $\bar{x}_B$ = sample mean of Group B
- s_A^2 = sample variance of Group A
- s_B^2 = sample variance of Group B
- n_A = sample size of Group A
- n_B = sample size of Group B

Degrees of freedom :

$$df = n_A + n_B - 2$$

Decision Rule :

If

$$|t| > t_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$$t_{critical} = t_{\frac{\alpha}{2},(n_A+n_B-2)} \text{ (i.e.,the tabulated value of t statistics)}$$

In [4]:
```
# Sample data
group_a = np.array([75, 78, 74, 77, 76, 73, 79, 80])  # Group A
group_b = np.array([72, 70, 68, 71, 69, 73, 67, 70])  # Group B

# t-test for independence
t2_stat, t2_pvalue = stats.ttest_ind(group_a, group_b, equal_var = True)

alpha = 0.05
```

```python
n_a = len(group_a)
n_b = len(group_b)
df_2 = n_a + n_b - 2

# two-tailed critical value
t2_critical = t.ppf(1 - alpha/2, df_2)

print("t-test for independence")
print("="*25)
print(f"Degree of freedom : {df_2}")
print(f"Critical t value : {t2_critical:.4f}")
print(f"Calculated t-statistic : {t2_stat:.4f}")
print(f"p-value : {t2_pvalue:.6f}")
print("\nDision :")
if t2_pvalue < 0.05 or abs(t2_stat) > t2_critical :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
t-test for independence
=========================
Degree of freedom : 14
Critical t value : 2.1448
Calculated t-statistic : 5.8138
p-value : 0.000045

Dision :
The Null Hypothesis is Rejected
```

Conclusion :

Here,

$$|t| = 5.8138 > 2.1448 = t_{critical}$$

or

$$p - value = 0.000045 < 0.05$$

So, we may **reject** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, there is a **significant difference** in the mean marks of students taught by the two teaching methods.

-: Methodology :-

Mann–Whitney U test : (Non-Parametric)

The **Mann–Whitney U test** (also called the Wilcoxon rank-sum test) is a non-parametric alternative to the independent (two-sample) t-test.

Step-by-step procedure

Step 1: State the hypothesis value

- **Null Hypothesis (H_0) :** $\mu_A = \mu_B$ (**i.e.,** There is no significant difference in the mean marks of students taught by the two methods.)
- **Alternative Hypothesis (H_1) :** $\mu_A \neq \mu_B$ (**i.e.,** There is a significant difference in the mean marks of students taught by the two methods.)

Step 2: Combine both samples

Step 3: Assign the rank to all observations

Step 4: Compute the sum of ranks for each group

i.e.,

- For Group A : R_A
- For Group B : R_B

Step 5: Compute the U statistic

Compute the U statistics for both Groups as :

- For Group A :

$$U_A = n_A n_B + \frac{n_A(n_A+1)}{2} - R_A$$

- For Group B :

$$U_B = n_A n_B + \frac{n_B(n_B+1)}{2} - R_B$$

So, the U statistic as combine is :

- $$U = min(U_A, U_B)$$

Step 6: Decision Rule : If

$$U > U_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$U_{critical} = U_{\alpha,n}$ (i.e.,the tabulated value of Mann–Whitney U test statistics, for sample s

In [5]:
```
# Mann-Whitney U test
u_stat, u_pvalue = stats.mannwhitneyu(group_b, group_a, alternative='two-sided')

print("Mann-Whitney U test")
```

```python
print("="*20)
print(f"U statistics : {u_stat:.4f}")
print(f"p-value : {u_pvalue:.6f}")
print("\nDision :")
if u_pvalue < 0.05 :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
Mann-Whitney U test
====================
U statistics : 0.5000
p-value : 0.001112

Dision :
The Null Hypothesis is Rejected
```

Conclusion :

Here,

$$p - value = 0.001112 < 0.05$$

So, we may **reject** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, there is a **significant difference** in the mean marks of students taught by the two teaching methods.

Case 3 :

Weights (kg) of individuals before and after a diet program:

Person	Before	After
1	72	70
2	75	70
3	68	62
4	80	77
5	77	75

Test whether the diet program is effective. (paired t-test & Wilcoxon Signed-Rank test)

-: Methodology :-

Paired t-test : (Parametric)

The **paired t-test** is used when we want to check whether the mean difference between two related (paired) observations is significantly different from zero.

Hypotheses :

- **Null Hypothesis (H_0) :** $\mu_d = 0$ (**i.e.,** There is no significant reduction in weight due to the diet program.)
- **Alternative Hypothesis (H_1) :** $\mu_d > 0$ (**i.e.,** The diet program reduces weight significantly.)

Test Statistic :

The paired t-test reduces to a one-sample t-test on the differences. And the *t* statistic is :

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

Where

- d = $x_{before} - x_{after}$ $(diffrence)$
- $\bar{d}$ = mean of differences
- s_d = standard deviation of differences
- n = number of pairs

Degrees of freedom :

$$df = n - 1$$

Decision Rule :

If

$$|t| > t_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$$t_{critical} = t_{\alpha,(n-1)} \text{ (i.e.,the tabulated value of t statistics)}$$

In [6]:
```
weights_before = np.array([72, 75, 68, 80, 77])
weights_after = np.array([70, 70, 62, 77, 75])

t3_stat, t3_pvalue = stats.ttest_rel(weights_before, weights_after)

alpha = 0.05
n = len(weights_before)
df_3 = n - 1
```

```python
# two-tailed critical value
t3_critical = t.ppf(1 - alpha, df_3)

print("Paired t-test")
print("="*20)
print(f"Degree of freedom : {df_3}")
print(f"Critical t value : {t3_critical:.4f}")
print(f"Calculated t-statistic : {t3_stat:.4f}")
print(f"p-value : {t3_pvalue:.6f}")
print("\nDision :")
if t3_pvalue < 0.05 or abs(t3_stat) > t3_critical :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
Paired t-test
====================
Degree of freedom : 4
Critical t value : 2.1318
Calculated t-statistic : 4.4313
p-value : 0.011411

Dision :
The Null Hypothesis is Rejected
```

Conclusion :

Here,

$$|t| = 4.4313 > 2.1318 = t_{critical}$$

or

$$p - value = 0.011411 < 0.05$$

So, we may **reject** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, the diet program is **effective in reducing weight.**

-: Methodology :-

Wilcoxon signed-rank test : (Non-Parametric)

The **Wilcoxon signed-rank test** is a non-parametric alternative to the paired t-test (and one-sample t-test). We use it when we cannot assume normality but still want to test a median difference.

Step-by-step procedure

Step 1: State the hypothesis value

- **Null Hypothesis (H_0) :** $\mu_d = 0$ (**i.e.,** There is no significant reduction in weight due to the diet program.)
- **Alternative Hypothesis (H_1) :** $\mu_d > 0$ (**i.e.,** The diet program reduces weight significantly.)
- **Step 2:** Compute differences For each observation i :

$$d_i = x_{before} - x_{after} \text{ (diffrence)}$$

Step 3: Remove zero differences

- Any $d_i = 0$ is dropped.
- Reduce sample size accordingly

Step 4: Take the absolute value of differences

$$|d_i|$$

Step 5: Rank the absolute differences

- Rank $|d_i|$ from smallest to largest.
- If ties occur then assign average ranks

Step 6: Assign signs to ranks (Give each rank the sign of its original difference)

- If $d_i > 0$ then rank is positive
- If $d_i < 0$ then rank is negative

Step 7: Compute the sum of ranks

- W^+ = sum of positive signed ranks
- W^- = sum of negative signed ranks

Step 8: Compute test Wilcoxon test statistic (W)

$$W = min(W^+, W^-) \text{ (i.e., the Wilcoxon signed-rank statistic)}$$

Step 6: Decision Rule :

If

$$W > W_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$$W_{critical} = W_{\alpha,n} \text{ (i.e.,the tabulated value of wilcoxon statistics W, for sample size n)}$$

In [7]:
```
# Wilcoxon signed-rank test
w2_stat, w2_pvalue = stats.wilcoxon(weights_before - weights_after)

print("Wilcoxon signed-rank test")
print("="*25)
print(f"Wilcoxon statistic : {w2_stat:.4f}")
print(f"p-value : {w2_pvalue:.4f}")
print("\nDision :")
if w2_pvalue < 0.05 :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
Wilcoxon signed-rank test
=========================
Wilcoxon statistic : 0.0000
p-value : 0.0625

Dision :
Failed to reject the Null Hypothesis
```

Conclusion :

Here,

$$p - value = 0.0625 > 0.05$$

So, we may **accept** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, the diet program is not effective.

Case 4 :

A survey was conducted on gender and preference for online courses.

Gender	Prefer Online	Prefer Offline
Male	30	20
Female	25	35

Test whether gender and course preference are independent.

-: Methodology :-

Chi-Square Test of Independence : (Parametric)

The **Chi-Square Test of Independence** is used to check whether two categorical variables are related (associated) or independent of each other.

Hypotheses :

- **Null Hypothesis (H_0) :** Gender and Course Preference are independent.
- **Alternative Hypothesis (H_1) :** Gender and Course Preference are not independent.

Test Statistic :

The Chi-square statistic is :

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Where

- O = observed frequency
- E = expected frequency
- and

$$E = \frac{(\text{Row total})(\text{Column total})}{\text{Grand total}}$$

Degrees of freedom :

$$df = (r-1)(c-1)$$

Where

- r = number of rows
- c = number of columns

Decision Rule :

If

$$\chi^2 > \chi^2_{critical}$$

or

$$p - value < \alpha$$

Then, reject the null hypothesis H_0 at α level of significance.

Where,

$$\chi^2_{critical} = \chi^2_{\alpha,d.f.} \text{ (i.e.,the tabulated value of Chi-square statistic)}$$

In [8]:
```
# Sample data
observed = np.array([[30,20],[25,35]])

#Chi-square test
chi2_stat, c_pvalue, df_4, expected = stats.chi2_contingency(observed)

alpha = 0.05
chi2_critical = chi2.ppf(1 - alpha, df_4)
```

```python
print("Chi-square test of independence")
print("="*35)
print("Degrees of Freedom :", df_4)
print("Expected Frequencies :\n", expected)
print("\nCalcuated Chi-square statistic :", round(chi2_stat,4))
print("Critical Chi-square value :", round(chi2_critical,4))
print("p-value :", round(c_pvalue,4))
print("\nDision :")
if c_pvalue < 0.05 or abs(chi2_stat) > chi2_critical :
    print("The Null Hypothesis is Rejected")
else :
    print(f"Failed to reject the Null Hypothesis")
```

```
Chi-square test of independence
===================================
Degrees of Freedom : 1
Expected Frequencies :
 [[25. 25.]
 [30. 30.]]

Calcuated Chi-square statistic : 2.97
Critical Chi-square value : 3.8415
p-value : 0.0848

Dision :
Failed to reject the Null Hypothesis
```

Conclusion :

Here,

$$\chi^2 = 2.97 < 3.8415 = \chi^2_{critical}$$

or

$$p - value = 0.0848 > 0.05$$

So, we may **accept** the **null hypothesis** (H_0) at 5 % level of significance.

Therefore, Gender and Course Preference are independent.
