PRACTICAL - 11

•) AIM : To find Spearmann's Rank Correlation.

•) EXPERIMENT :

X	Y
10	8
12	9
14	11
15	14
18	16
20	19
22	21

•) THEORY :

(•) $R(x_i) =$ rank of x_i

(•) $R(y_i) =$ rank of y_i

(•) $d_i = R(x_i) - R(y_i)$

(•) $\rho = 1 - \frac{6 \Sigma d_i^2}{n(n^2-1)}$

•) CALCULATIONS :

(•) $n = 7$

(•) $d_i^2 = 0$

(•) $\rho = 1$

•) RESULT : The Spearman's Rank Correlation for the given data set is 1.

PRACTICAL - 10

•) AIM : To apply Kendall's Tau Coefficient.

•) EXPERIMENT :

X	Y
1	2
2	3
3	1
4	4
5	5
6	6

•) THEORY :

(•) Concordant pairs (C) if $(x_i - x_j)(y_i - y_j) > 0$

(•) Discordant pairs (D) if $(x_i - x_j)(y_i - y_j) < 0$

(•) Total no. of possible pairs $= {}^nC_2$

(•) $\tau = \dfrac{C - D}{{}^nC_2}$.

•) CALCULATIONS :

(•) Concordant = 13

(•) Discordant = 2

(•) $n = 6$

(•) Pairs = 15

(•) $\tau = 0.73$

•) RESULT : The Kendall's Tau Coefficient, $\tau = 0.73$

PRACTICAL - 9

- AIM : To apply Kruskal-Wallis one-way ANOVA.
- EXPERIMENT :

Group 1	Group 2	Group 3
12	10	15
15	12	17
14	14	18
13	13	16
16	11	19

- THEORY :

(•) $R_i = \sum_{j=1}^{n_i} R_{ij}$

(•) $H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$

(•) Reject H_0 if : $H > \chi^2_{critical}(k-1, \alpha)$

- CALCULATIONS :

(•) $N = 15$

(•) $H = 5.93$

(•) p-value $= 0.05156$

- Result : Since at 5% level of significance p-value > 0.05, therefore we reject H_0

PRACTICAL - 8

- AIM : To apply Wilcoxon Rank-Sum test
- EXPERIMENT :

Group 1	Group 2
45	40
47	42
49	44
50	46
52	48
54	51
56	53
58	55

- THEORY :

(•) $\mu_w = \frac{n_1(N+1)}{2}$

(•) $\sigma_w = \sqrt{\frac{n_1 n_2 (N+1)}{12}}$

(•) $Z = \frac{W_1 - \mu_w}{\sigma_w}$

- CALCULATIONS :

(•) $G_1 = 8$

(•) $G_2 = 8$

(•) $Z = 1.47029$

- RESULT : Since the value of $Z = 1.47029$ i.e, $Z_{cal} < Z_{tab}$, therefore we accept the null hypothesis H_0.

- CALCULATIONS :

⇒ Median Test -

$\overline{X} = 49.5$

$\chi^2 = 1$

$Z = 1$

⇒ Mann - Whitney U test. -

$G1 = 8$

$G2 = 8$

$U = 18$

$Z = -1.4703$

- RESULT :

⇒ Since in Median test $Z = 1$

$Z_{cal} < Z_{tab}$, therefore we accept H_0.

⇒ Since in Mann - Whitney U Test

$Z = -1.4703$

$Z_{cal} < Z_{tab}$, therefore we accept H_0.

PRACTICAL - 7

- **AIM** : To apply Median Test and Mann-Whitney U test

- **EXPERIMENT** :

Group 1	Group 2
45	40
47	42
49	44
50	46
52	48
54	51
56	53
58	55

- **THEORY** :

⇒ Median Test :

- $\bar{X}$ = Median of all observations
- $N = \sum_{i=1}^{k} (a_i + b_i)$
- $\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{2} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$

where O_{ij} = observed frequency in cell (i,j)

$E_{ij} = \frac{(\text{row total}) \times (\text{column total})}{N}$

⇒ Mann-Whitney U-test :

- $\mu_U = \frac{n_1 n_2}{2}$
- $\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$
- $Z = \frac{U - \mu_U}{\sigma_U}$

PRACTICAL - 6

- AIM : To apply Kolmogorov - Smirov One sample Test.

- EXPERIMENT :

Data

2.1
2.4
3.0
3.2
3.3
3.5
3.8
4.0
4.2
4.4

- THEORY :

$F_n(x)$ = emperical cdf

$F_0(x)$ = theoretical cdf under H_0

$D_n = \sup |F_n(x) - F_0(x)|$

- CALCULATIONS :

As shown in Excel.

$D_n(x) = 0.1064418$.

- RESULT :

Since the value of D at 0.05 is 0.409

$D_n(x) < D_{tab}$

∴ Accept the null hypothesis H_0.

PRACTICAL - 5

- AIM : To apply Wald - Wolfowitz Runs test.
- EXPERIMENT :

Value	Group
10	A
12	A
9	A
15	B
14	B
13	A
16	B
18	B
11	A
17	B

Use Wald - Wolfowitz Runs test to test the null hypothesis Ho. ~~Ho~~

- THEORY :

(•) $\mu_k = 1 + \frac{2n_1n_2}{n_1+n_2}$

(•) $\sigma_k^2 = \frac{2n_1n_2(2n_1n_2 - n_1 + n_2)}{(n_1+n_2)^2(n_1+n_2-1)}$

(•) $Z = \frac{k - \mu_k + 1/2}{\sigma_k}$

- CALCULATIONS :
  - (•) $U = 2$
  - (•) $n_1 = 5$
  - (•) $n_2 = 5$
  - (•) $\mu_k = 6$
  - (•) $\sigma_k = 1.66$
  - (•) $Z = -2.1$

- RESULT : Since $|-2.1| > |-1.96|$ i.e, $Z_{cal} > Z_{tab}$

∴ We reject the null hypothesis

PRACTICAL - 4

- **AIM :** To apply Wilcoxon Signed-Rank Test.
- **EXPERIMENT :**

Before	After
45	50
52	47
56	56
48	52
49	51
50	49
55	58
60	61
53	55
51	50

Use Wilcoxon Signed-Rank Test.

- **THEORY :**

(•) $T = \min(T^+, T^-)$

(•) $\text{Mean} = \frac{n(n+1)}{4}$

(•) $\text{Variance} = \frac{n(n+1)(2n+1)}{24}$

(•) $$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

- **CALCULATIONS :**

(•) $T^+ = 41.5$ (•) $T^- = 12.5$

(•) $n = 9$ (•) $T = 12.5$

(•) $\mu = 22.5$ (•) $\sigma_{\overline{z}} = 8.44$

(•) $Z = -1.18$

- **RESULT :** Since $Z_{cal} < Z_{tab}$, we accept the null hypothesis H_0.

PRACTICAL - 3

- **AIM :** To determine if there is significant difference b/w Before and After measurements using the Sign test.

- **EXPERIMENT :**

Before	After
45	50
52	47
56	56
48	52
49	51
50	49
55	58
60	61
53	55
51	50

Use sign test to determine if there is significant difference b/w before and after.

- **THEORY :**

(•) $\mu = np$

(•) $\sigma^2 = npq$

(•) $Z = \frac{U - np}{\sqrt{npq}}$

- **CALCULATIONS :**

(•) $\mu = 4.5$
(•) $\sigma^2 = 2.25$
(•) $n = 9$
(•) ~~Z = 1~~ $U = 6$
(•) $Z = 1$

- **RESULT :** We fail to reject H_0 since $Z_{cal} < Z_{tab}$ at 5% level of significance.

i.e, there is no significant difference b/w before and after measurement.

PRACTICAL - 2

- **AIM :** To test for randomness using Run test.
- **EXPERIMENT :** For the given data set
  H, T, H, T, T, H, H, T, H, T test for the randomness using Run test.
- **THEORY :**
  - Mean(μ)= $\frac{2n_1 n_2}{(n_1+n_2)} + 1$
  - Variance(σ^2) = $\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1+n_2)^2 (n_1+n_2-1)}$
  - $Z = \frac{U - \mu}{\sigma}$

    where U = no. of runs
    n_1 = no. of heads
    n_2 = no. of tails.
- **CALCULATIONS :**
  - $\mu = 6$
  - $\sigma^2 = 2.22$
  - $Z = \frac{8-6}{\sqrt{2.22}} = 1.34$
- **RESULT :**

  At 5% level of significance
  $Z_{cal} < Z_{tab}$ i.e $1.34 < 1.96$
  $\therefore$ Accept H_0. i.e, the data set is random.

PRACTICAL - 1

AIM : To estimate the value of Quantile and Emperical distribution.

EXPERIMENT : For the data set :

5, 6, 7, 8, 9, 10, 11, 12, 13, 14 find the quantile & emperical distribution.

THEORY :

- Quantile – The p^{th} quantile for the 100 p^{th} percentile is that value of the random variable X, say X_p, such that 100 p % of the values of X in the population are less than or equal to X_p.
- Emperical function – Let $x_1, x_2, \ldots, x_n$ be a random sample from cdf $F(x)$. Let $x_{(1)}, x_{(2)}, \ldots x_{(n)}$ is corresponding order statistic. The emperical dist. f/n :

$$F_n(x) = \begin{cases} 0 & \text{if } x < x_{(1)} \\ j/n & \text{if } x_{(j)} \leq x \leq x_{(j+1)}, j = 1, \ldots, n \\ 1 & \text{if } x \geq x_{(n)} \end{cases}$$

CALCULATIONS :

$Q_1 = 7.25$ $Q_2 = 9.5$ $Q_3 = 11.75$

Edf as shown in the excel file.

RESULT : The edf shows how the data accumulate from the smallest to the largest value. It increases by 0.1 at each point because there are 10 observations.

9.5 is the median of the given data set.