Wilcoxon signed rank test (Single sample case)

The titanium content in an aircraft-grade alloy is an important determinant of strength. A sample of 20 test coupons reveals the following titanium content (in %)

8.32	8.05	8.93	8.65	8.25	8.46	8.52	8.35	8.36	8.41
8.42	8.30	8.71	8.75	8.60	8.83	8.50	8.38	8.29	8.46

The **median** titanium content should be 8.5 %. Use the **Wilcoxon signed rank** to investigate this hypothesis and $\alpha = 0.05$.

Solution:

Random variable X = Titanium content in an aircraft-grade alloy (%)

Sample size (n) = 20 (small sample case)

Step 1: Null and Alternative hypothesis

Let M be the titanium content in an aircraft-grade allow

H_0: M = 8.5 (The median titanium content is 8.5 %)

H_1: M ≠ 8.5 (The median titanium content is not 8.5 %)

Step 2: Level of significance for the test

The chosen level of significance for the test is 5 % = 0.05

Step 3: Test Statistic

Since n = 20 (< 25), we test the hypothesis using small sample approach.

The steps in finding test statistic is given by,

1. Compute the difference $D_i = X_i - M_0$ in case of single sample.
2. Find the absolute value of each difference D_i , i = 1, 2, …, n.
3. Rank each absolute value from lowest to highest. In case of ties, assign the average rank.
4. Give each rank a plus (+) or minus (-) sign according to sign in the difference column.
5. Find the sum of positive ranks and the sum of negative ranks. Let T^+ be the sum of positive ranks and T^- be the sum of negative ranks.

Test Statistic (T) = Min $(T^+, |T^-|)$

Step 4: Calculated W

Titanium Content	Diff D = X - 8.5	Abs D	Rank of abs D	Signed Rank
8.32	-0.18	0.18	11	-11
8.05	-0.45	0.45	19	-19
8.93	0.43	0.43	18	18
8.65	0.15	0.15	9.5	9.5
8.25	-0.25	0.25	15.5	-15.5
8.46	-0.04	0.04	2.5	-2.5
8.52	0.02	0.02	1	1
8.35	-0.15	0.15	9.5	-9.5
8.36	-0.14	0.14	8	-8
8.41	-0.09	0.09	5	-5
8.42	-0.08	0.08	4	-4
8.3	-0.2	0.2	12	-12
8.71	0.21	0.21	13.5	13.5
8.75	0.25	0.25	15.5	15.5
8.6	0.1	0.1	6	6
8.83	0.33	0.33	17	17
8.5	0	omit	#VALUE!	
8.38	-0.12	0.12	7	-7
8.29	-0.21	0.21	13.5	-13.5
8.46	-0.04	0.04	2.5	-2.5

Sum of positive rank (T^+) = 80.5

Sum of negative rank ($|T^-|$) = - 109.5

Calculated T = Min (T^+, $|T^-|$) = (80.5, 109.5) = 80.5

Step 5: Tabulated or Critical T

The test is two sided and level of significance is 5 %.

The reduced sample size (n) = 20 -1 = 19

The critical T from table is given by,

$T_{0.025}(19) = 46$

AR: T > 46

RR: T ≤ 46

Step 6: Statistical Decision

Since cal T = 80.5 > Critical T = 46, we accept our null hypothesis at 5 % level of significance.

Step 7: Conclusion

The median Titanium content in the ore is 8.5 % or higher.

Problem 2:

Two models of machines are under consideration for purchase. An organization has one of each type for trial. Each operator, out of the team of 25 operates uses each machine for a fixed length of time. Their outputs are:

Operator	Machine I	Machine II	Operator	Machine I	Machine II
1	82	80	14	65	60
2	68	71	15	70	73
3	53	46	16	55	48
4	75	58	17	75	58
5	78	60	18	64	60
6	86	72	19	72	76
7	64	38	20	55	60
8	54	60	21	70	50
9	62	65	22	45	30
10	70	64	23	64	30
11	51	38	24	58	55
12	80	79	25	65	60
13	64	37			

Is there any significant **difference** between the output capacities of two machines? Use $\alpha = 5\%$.

Solution:

Let X_1 and X_2 be the number of outputs using machine I and II respectively.

Step 1: Null and alternative hypothesis

Let M1 and M2 be the median no. of outputs using machine I and machine II respectively.

Null Hypothesis

H0: $M_1 = M_2$ (There is no significant difference in the median no. of output using machine I and II respectively i.e., output capabilities of two machines are same)

H1: $M1 \neq M2$ (There is significant difference in the median no. of output using machine I and II respectively i.e., output capabilities of two machines are not same)

Step 2: Level of significance for the test

The given level of significance for the test = 5 %

Step 3: Test Statistic

Since n = 25, we will solve this problem using large sample approximation (i.e. using z distribution)

The steps in finding the Z statistic is given by,

1. Compute the difference $D_i = X_{1i} - X_{2i}$ = Before – After, where X_1 is variable before treatment and X_2 is variable after treatment.
2. Find the absolute value of each difference D_i , i = 1, 2, …, n.
3. Rank each absolute value from lowest to highest. In case of ties, assign the average rank.
4. Give each rank a plus (+) or minus (-) sign according to sign in the difference column.
5. Find the sum of positive ranks and the sum of negative ranks. Let T^+ be the sum of positive ranks and T^- be the sum of negative ranks.

Now,

The mean of T (μ_T) = $\dfrac{n(n+1)}{4}$

The standard deviation of T (σ_T) = $\sqrt{\dfrac{n(n+1)(2n+1)}{24}}$

Thus Z equivalent of W values is given by,

$$Z_T = \frac{T-\mu_T}{\sigma_T}$$

Where, W is the smaller sum in absolute value of signed ranks and n is the number of observations or number of paired observations where difference is not zero.

Step 4: Calculated Z

Machine I	Machine II	Diff Di	Abs Di	Rank of Di	Signed Rank
82	80	2	2	2	2
68	71	-3	3	4.5	-4.5
53	46	7	7	14.5	14.5
75	58	17	17	19.5	19.5
78	60	18	18	21	21
86	72	14	14	17	17
64	38	26	26	23	23
54	60	-6	6	12.5	-12.5

62	65	-3	3	4.5	-4.5
70	64	6	6	12.5	12.5
51	38	13	13	16	16
80	79	1	1	1	1
64	37	27	27	24	24
65	60	5	5	10	10
70	73	-3	3	4.5	-4.5
55	48	7	7	14.5	14.5
75	58	17	17	19.5	19.5
64	60	4	4	7.5	7.5
72	76	-4	4	7.5	-7.5
55	60	-5	5	10	-10
70	50	20	20	22	22
45	30	15	15	18	18
64	30	34	34	25	25
58	55	3	3	4.5	4.5
65	60	5	5	10	10

sum of positive rank (T+) = 281.5

sum of negative rank (T-) = - 43.5

Abs T- = 43.5

Now,

$$\mu_T = \frac{n(n+1)}{4} = 25 * 26 / 4 = 162.5$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{25(25+1)(2x25+1)}{24}} = 37.165$$

finally

$$Z_T = \frac{T-\mu_T}{\sigma_T} = \frac{43.5-162.5}{37.16} = -3.201$$

Step 5: Tabulated Z

Test is two sided and level of significance is 5%

Critical Z is given by,

Zc = 1.96

Acceptance Region: - 1.96 < Z < + 1.96

Rejection Region: Z ≤ - 1.96 OR Z ≥ + 1.96

Step 6: Statistical Decision

Since cal Z = - 3.2 < critical Z = - 1.96, we reject our null hypothesis at 5 % level of significance

Step 7: Conclusion

There is significant difference between median no. of outputs by using machine I and machine II i.e. output capacities of machine I and machine II are not same)

Since the median output of machine I i.e., M1 = 65 and median output of machine II i.e. M2 = 60, we can say output capacity of machine I is higher than machine II.

Small sample approach

Since the test is two sided the appropriated test statistic is given by

The test statistic = Min (T+, $|T^-|$)

Calculated W

Cal T = min (T+, $|T^-|$) = min (43.5, 181.5) = 43.5

Tabulated W

Test is two sided and level of significance is 5 %

From table the critical T is given by,

Tc = $W_{0.025}(25)$ = 89

AR: T > 89

RR: T ≤ 89

Statistical Decision

Since, cal T = 43.5 falls in the rejection region T ≤ 89, we reject our null hypothesis at 5 % level of significance