Date 22.04.22

F.P&S

Further Probability and Statistics

Content. Page.

1. The sign test 1
2. Wilcoxon signed-rank test 2-4
3. Wilcoxon matched-pairs signed-rank test 5-7
4. Wilcoxon rank-sum test. 8-10

Non-Parametric Tests
Notes and Revision.

SP-20 / S-20 / W-20 / S-21 / W-21

Suresh Goel
(Former Director)
Alliance World School
Noida - Delhi - NCR
INDIA.

(+91 9810444804)

F. P&S

Non-Parametric Tests

§ Non-Parametric tests:

When samples are small and nothing can be assumed about the underlying population distribution, we use non-parametic testing.

§ The sign test:

In sign test the null hypothesis is that the population **median** is a specific value is true. The test works by considering the differences in the sample data values from the hypothesised median.

§ **Wilcoxon signed-rank test:**

The only assumption required is that, under the null hypothesis, the data is symmetrically distributed. (i) We find the deviation of each data from median.

(ii) Then find the absolute value of these deviations.

(iii) These absolute values are then ranked. (1 to n)

Smallest deviation is ranked first. -- n the last rank.

(iv) These deviations are then given signs.

(v) The sum of all positive values gives a value P. and the absolute value of the sum of negative Ranks gives value Q.

The sum of $P+Q = \frac{1}{2}n(n+1)$ [n is the number of data)

Now data being symmetrically distributed about the median so the values of P and Q be close to $\frac{1}{2}(P+Q)$.

We crate a test statistic 'T' which is lower of P and Q.

{ This can be compared to table - "Wilcoxon signed-rank Test"
Note: At 5% significan level for one tailed test and let $n=12$, $T \leq 17$

Compare it with lower of P and Q

§ Wilcoxon matched-pairs signed-ranks test.

§ Wilcoxon rank-sum test.

1(a) State briefly the circumstances under which a non-parametric test of significance should be used rather than a parametric test --[1].

The level of pollution in a river was measured at 12 randomly chosen locations. The results, in suitable units, are shown below, where higher values represent greater pollution.

5.62, 5.73, 6.55, 6.81, 6.10, 5.75, 5.87, 6.47, 5.86, 6.26, 6.99, 5.91

(b) Use a Wilcoxon signed-rank test whether the average pollution level in the river is more than 6.00. Use 5% significance level. --[6]

SP-2004/Q1

Solution (a) When the population cannot be assumed to be normally distributed.

(b) H_0: Population median is 6.00; H_1: Population median is greater than 6.00.

Now deviation of data from the median 6.00 and signed ranks; $n = 12$

Deviation	−.38	−0.27	0.55	0.81	0.10	−0.25	−0.13	0.47	−0.14	0.26	0.99	−0.09
Abs. devi	0.38	0.27	0.55	0.81	0.10	0.25	0.13	0.47	0.14	0.26	0.99	0.09
Rank	8	7	10	11	2	5	3	9	4	6	12	1
Signed Rank	−8	−7	10	11	2	−5	−3	9	−4	6	12	−1

Sum of deviations (ranks) of all positive $P = 12+11+10+9+6+2 = 50$

Sum of deviations (ranks) of all negative $Q = 8+7+5+4+3+1 = 28$

Test statistic (lower of P & Q is Q) $T = 28$ ✓

Now $\frac{1}{2}(P+Q) = \frac{1}{2}(50+28) = 39$,

at 5% significance value (for one tailed), $0.05 \rightarrow T \leq 17$ (use the table for Wilcoxon signed-rank test) and $n = 12$, critical region $28 > 17$, does not lie in the critical region.

$\therefore$ H_0 is accepted

or insufficient evidence that the median is greater than 6.00.

2. The times, in milliseconds, taken by a computer to perform a certain task were recorded on 10 randomly chosen occasions.

6·44 6.16 5.62 5.82 6.51 6.52 6.19 6.42 6.34 6.28

It is claimed that the median time to complete the tas is 6·4 milli sec.

(a) Carray out a Wilcoxon signed-rank test at the 5% significance level to test this claim. ---[6]

(b) State an underlying assumption that is made when using a Wilcoxon Signed-rank test. [S-20/41/Q2] --[1]

Solution: Ho: Population median is 6·40; population median ≠ 6·40

Now to to calculate the deviation from median 6·40 & signed ranks.

Deviation;	0·04	−0·24	−0·78	−0.58	0·11	0·22	−0·21	0·02	−0·06	−0·12
Abs. Devi.	0·04	0·24	0·78	0·58	0·11	0·22	0·21	0·02	0·06	0·12
Ranks	2	8	10	9	4	7	6	1	3	5
Signed Ranks	2	−8	−10	−9	4	7	−6	1	−3	−5

Sum of all positive ranks $P = 4+7+2+1 = 14$; $(n=10)$ $\{P+Q = 55 = \frac{1}{2}10\times(10+1)$

Sum of all negative ranks $Q = -8-10-9-6-3-5 = -41$

Test Statistic (lower of P and Q) $T = 14$.

At 5% Significance level (for two tailed) at 0·025 and $n=10$; $T=8$

Critical region $T \leq 8$.

$14 > 8 \Rightarrow$ Ho is accepted.

Hence insufficient evidence to rejection the claim.

3. Metal rods produced by a certain factory are claimed to have a median breaking strength of 200 tonnes. For a random sample of 9 rods, the breaking strengths, measured in tonnes, were as follows:

210 186 188 208 184 191 215 198 196

A scientist believes that the median strength of the metal rods produced by this factory is less than 200 tonnes.

(a) Use a wilcoxon signed-rank test, the 5% significance level, to test whether there is evidence to support the scientist's belief. ---[6]

(b) Give a reason why a Wilcoxon signed-rank test is preferable to a sign test, when both are valid. [W-20/41/Q2] ---[1]

Solution: H_0: median = 200 ; H_1: median < 200

Deviation from Med 200	10	−14	−12	8	−16	−9	15	−2	−4
Abso. deviation	10	14	12	8	16	9	15	2	4
Rank	5	7	6	3	9	4	8	1	2
Signed Rank	5	−7	−6	3	−9	−4	8	−1	−2

Sum of ranks (all positive) $P = 5 + 3 + 8 = 16$

Sum of ranks (all negative) $Q = 7 + 6 + 9 + 4 + 1 + = 29$

$\left[\begin{array}{l} P+Q = 16+29 = 45 \\ n = 9;\ P+Q = \frac{1}{2} \cdot n(n+1) \end{array}\right]$

Test statistic (lower of P & Q) $T = 16$ ✓

Now at 5% significance level (for one tailed) and $n = 9$.
$T \leq 8$ critical region (Wilcoxon Sign-rank test table).

Here $16 > 8 \rightarrow H_0$ is is accepted.

Insufficient evidence to support scients belief.

Wilcoxon matched-pairs signed-rank test

4. A large school is holding an essay completition and each student has submitted an essay. To ensure fairness, each essay is given a mark out of 100 by two different judges. The marks awarded to the essay submitted by a random sample of 12 students are shown in the following table:

Student	A	B	C	D	E	F	G	H	I	J	K	L
Judge 1	62	74	52	48	68	55	56	64	37	70	81	59
Judge 2	65	70	47	49	76	74	67	54	50	77	72	75

(a) One of the students claim that Judge 2 is awarding higher marks than Judge 1. Carry out a "Wilcoxon matched-pairs signed-rank test" at 5% significance level to test whether the data supports the student's claim. --[7]

It is observed later the marks awarded to Student A have been entered incorrectly. In fact Judge 1 awarded 65 marks and Judge 2 awarded 62 marks.

(b) By considering how this change affects the test statistic, explain why the conclusion of the test carried out in part (a) remain the same. [W-20/42/Q2] --[2]

Solution: H_0: difference in (population) median $= 0$; H_1: diff in median < 0 (or > 0)

Differences	−3	4	5	−1	−8	−19	−11	10	−13	−7	9	−16
Abs. Diff	3	4	5	1	8	19	11	10	13	7	9	16
Rank	2	3	4	1	6	12	9	8	10	5	7	11
Signed Rank	−2	3	4	−1	−6	−12	−9	8	−10	−5	7	−11

Sum of difference (Ranks) all positive $P = 3+4+8+7 = 22$ $(P+Q = 78)$

Sum of diff. (Ranks) all Negative $Q = 2+1+6+12+9+10+5+11 = 56$ $(n = 12)$

Test Statistic (lower of P and Q is P) $\rightarrow T = 22$ and $n = 12$

at 5% Significance level for one tailed test for $n = 12$ (Wilcoxon Signed-Rank Test)

$T \leq 17$ ✓ ; here $22 > 17 \rightarrow H_0$ is accepted.

data does not support students claim.

(b) Rank for A becomes $+2$; Change of sign difference can only reduced evidence in favour of the claim, (Still > 17)

5. Georgio has designed two new uniforms X and Y for employees of an airline company. A random sample of 11 employees are each asked to assess each of the two uniforms for praticality and and appearance, and to give a total score out of 100.

Employee	A	B	C	D	E	F	G	H	I	J	K
Uniform X	82	74	42	59	60	73	94	98	62	36	50
Uniform Y	78	75	63	56	67	82	99	90	72	48	61

(a) Give a reason why a Wilcoxon signed-rank test may be more appropriate than a t-test for investigating whether there is any evidence of a preference for one of the uniforms. ---[1]

(b) Carry out a Wilcoxon match-pairs signed-rank test at the 10% significance level. $\boxed{S\text{-}21/41/Q5}$ [7]

Solution: (a) underlying distribution or population of differences is unknown, not known to be normal.

(b) H_0: difference of population medians $= 0$

H_1: difference of population median $\neq 0$

Differences	4	−1	−21	3	−7	−9	−5	8	−10	−12	−11
Abs. Diff.	4	1	21	3	7	9	5	8	10	12	11
Rank	3	1	11	2	5	7	4	6	8	10	9
Signed Rank.	3	−1	−11	2	−5	−7	−4	6	−8	−10	−9

Sum of Diff. (ranks) of all positive $P = 3+2+6 = 11$

Sum of diff (ranks) of all negative $Q = 1+11+5+7+4+8+10+9 = 55$

Test statistic (lower of P and Q is P) $T = 11$ ✓

At 10% significance level (Two tailed), 0.1 and $n = 11 \Rightarrow T = 13$.

Critical region $T \leq 13$

Now test statistic $11 < 13 \Rightarrow$ Reject H_0.

Sufficient evidence of a preference for one of the uniforms.

6. A company is developing a new flavour of chocolate by varying the quantities of ingredients. A random selection of 9 flavours of chocholate are judged by two tasters who each give marks out of 100 to each flavour of chocolate:

Chocolate	A	B	C	D	E	F	G	H	I
Taster 1	72	86	75	92	98	79	87	60	62
Taster 2.	84	72	74	95	85	87	82	75	68

Carry out a Wilcoxon matched-pair Signed-rank test at the 10% significance level to investigate whether, on average, there is a difference between marks awarded by the two tastors. --[7]

S-21/43/Q2

Solution: H_0: Difference in population medians $= 0$ ($m_a = m_b$)

H_1: Difference in population medians $\neq 0$ ($m_a \neq m_b$)

Differences	12	−14	−1	3	−13	8	−5	15	6
Abso. differences.	12	14	1	3	13	8	5	15	6
Ranks	6	8	1	2	7	5	3	9	4
Signed Ranks.	6	−8	−1	2	−7	5	−3	9	4

Sum of differences (Ranks) all positive $P = 6+2+5+9+4 = 25$

Sum of differences (Ranks) all negative $Q = 8+1+7+3 = 19$ (Smaller sum)

Test Statistic (Smaller of P and Q) → $T = 19$. ✓ ;

Now at 10% significance level for two tailed test; $n=9$,

Using (Wilcoxon - Signed-Rank test) $T \leq 8$ is the critical region.

$19 > 8$ → H_0 is accepted.

Hence there is insufficient evidence that marks differ.

Scanned with CamScanner

Wilcoxon rank-sum test

To carry out a Wilcoxon rank-sum test:

If the sample $(n+m) > 20$ use Normal distribution
Mean $\mu = \frac{1}{2} m(n+m+1)$
Variance $\sigma^2 = \frac{1}{12} nm(n+m+1)$

1. Rank all values from both values.
2. Find R_m, the sum of the ranks of the data points in sample size m. $m < n$
3. The test statistic W is the smaller of R_m and $m(n+m+1) - R_m$
4. Use tables to compare to the sampling distribution, rejecting H_0 if W lies in the critical region.

7. A biologists is studing the effect of nutrients on the heights to which plants grow. A random sample of 24 similar young plants is divided into two equal groups A and B. The plants in group A are fed with nutrients and water and the plants in group B are given only water. After four weeks, the heights, in cm, of each plant is measured as follows:

Group A	12.3	11.8	12.1	13.2	11.1	10.6	13.8	12.0	12.2	12.4	13.5	13.9
Group B	11.7	10.8	10.9	11.3	11.2	12.6	11.0	10.5	11.9	12.5	10.7	11.6

The biologist decides to carry out a test at the 5% significance level to test whether the nutrients have resulted in an increase in growth.

(a) She carries out a Wilcoxon rank-sum test. Give a reason why this is an appropriate choice of test. --[1]

(b) Carry out the Wilcoxon rank-sum test for these results. --[10]

[S-20/43/Q6]

Solution(a) Difference of location test for population not know to be normal.

(b) $H_0: m_x = m_y$; $H_1: m_x > m_y$; mean $= \frac{1}{2} m(n+m+1) = \frac{1}{2} \times 12 \times 25 = 150$

as sample $24 > 20 \rightarrow$ apply Normal distribution. Variance $= \frac{1}{12} \times nm(n+m+1) = \frac{12 \times 12}{12}(25) = 300$

GrpA, x	10.6	11.1	11.8	12.0	12.1	12.2	12.3	12.4	13.2	13.5	13.8	13.9
Rank, x	2	7	12	14	15	16	17	18	21	22	23	24
GrpB, y	10.5	10.7	10.8	10.9	11.0	11.2	11.3	11.6	11.7	11.9	12.5	12.6
Rank y	1	3	4	5	6	8	9	10	11	13	19	20

Total Ranks $R_m = 109$ and $R_n = 191$

Using Normal dis. $P(X \leq 109) = P\left(Z < \frac{109.5 - 105}{\sqrt{300}}\right)$
$= P(Z < -2.338)$

$= 1 - \phi(2.338) = 1 - 0.9903$

$p = 0.0097$, Now Significance level 5%

$0.0097 < 0.05 \rightarrow$ Reject H_0.

There is evidence of an effect on growth.

8. The blood cholesterol levels, measured in suitable units, of a random sample of 11 women and a random sample of 12 men are as:

Women	51	55	242	167	152	256	75	137	98	238	235	—
Men	311	262	170	302	175	320	220	260	72	351	86	333

Carry out a Wilcoxon rank-sum test, at the 5% significance level, to test whether, on average, there is a difference in Cholesterol levels between women and men. [W-21/41/Q6] -- [9]

Solution:

Women x	51	55	75	98	137	152	167	235	238	242	256	—
Rank x	1	2	4	6	7	8	9	13	14	15	16	
Men y	72	86	170	175	220	260	262	302	311	320	333	351
Rank y	3	5	10	11	12	17	18	19	20	21	22	23

$H_0: m_x = m_y$ and $H_1: m_x \neq m_y$ $\quad (m = 11,\ n = 12)$ $\quad \boxed{m+n > 20}$

using normal approximation with attempts at mean and variance.

Mean $= \frac{1}{2} m(n+m+1) = \frac{1}{2} \times 11 \times (12+11+1) = \frac{1}{2} \times 11 \times 24 = 132$ ✓

Variance $= \frac{1}{12} n \cdot m(n+m+1) = \frac{1}{12} \times 12 \times 11 \times 24 = 264$ ✓ ; $(R_m = \Sigma r_x$

using normal distribution (Total Rank $R_m = 95$)

Total Rank $R_n = 181$

$$P(X \leq 95) = P\left(Z < \frac{95.5 - 132}{\sqrt{264}}\right) = -$$

$$= P(Z < -2.246) = 1 - \phi(2.246) = 1 - 0.9877$$

Test Statistic $\quad p = 0.0123$

Now at 5% significance level for two tailed test;

Critical value $= 0.025$

here $0.0123 < 0.025 \Rightarrow$ Reject H_0.

Hence there is sufficient evidence of difference in levels.

9. Applicants for a particular college take a written test when they attend for interview. There are two different tests, A and B and each applicant takes one or the other. The interviewer wants to determine whether the medians of the distribution of marks obtained in the two tests are equal. The marks obtained by a random sample of 8 applicants who took test A and a random sample of 8 applicants who took test B are as follows:

Test A	46	32	29	12	33	18	25	40
Test B	36	28	49	37	48	35	41	31

(a) Carry out a Wilcoxon rank-sum test at 5% significance level to determine whether there is a difference in the population median marks obtained in the two test. ...[6]

The interviewer considers using the given information to carry out a paired sample t-test to determine whether there is a difference in the population means for the two tests.

(b) Give two reasons why it is not appropriate to use this test. ...[2]

W-21/42/Q4

Solution:
(a)

Test A - x	12	18	25	29	32	33	40	46
Rank x_A	1	2	3	5	7	8	12	14
Test B - y	28	31	35	36	37	41	48	49
Rank y_B	4	6	9	10	11	13	15	16

Test statistic $\sum^{Rank} x_A = R_m = 52$ ✓ ; $m = n = 8$

$H_0: m_x = m_y$ and $H_1: m_x \neq m_y$

$W =$ Smaller of R_m & $m(n+m+1) - R_m$ $\left\{ \begin{array}{l} m(n+m+1) - R_m \\ = 8 \times 17 - 52 = 84 \end{array} \right.$

$W = 52$ ✓

Now at 5% significance level (Two tailed), $m = n = 8$, $(p = 0.05)$
(using Wilcoxon Rank-sum table) → Critical value $= 49$ ✓

Here $52 > 49$, Accept H_0.

Insufficient evidence of difference in medians.

(b) Not a paired sample.
Underlying distribution is not normal.
(or Underlying population is unknown)