

Lecture 55: Paired signed-rank test with ties; Hodges–Lehmann estimator

Statistical Inference–II ■ Units IV–V: Nonparametric Inference

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

Previous lecture

Lecture 54: Wilcoxon signed-rank test: null distribution.

Today

- Wilcoxon Signed-Rank Test
 - Paired-Sample Form
 - Hodges–Lehmann Estimator
 - Sign Test versus Signed-Rank Test

For paired observations (X_i, Y_i) , use

$$D_i = Y_i - X_i$$

and apply the one-sample procedure.

Under a symmetric distribution of paired differences, the null hypothesis is that the location shift is zero.

Example: paired signed-rank test with ties

Example 55.1 (Paired signed-rank test with ties)

Apply the Wilcoxon signed-rank test to the blood-pressure data of the example “Paired sign test”, assuming that the distribution of the reduction D is symmetric.

Solution

Test H_0 : the distribution of D is symmetric about 0, against H_1 : it is symmetric about a positive value. Delete the zero difference, so $n = 11$, and rank the absolute differences, giving average ranks to ties:

D	-2	-3	3	5	5	6	7	8	9	10	12
$ D $	2	3	3	5	5	6	7	8	9	10	12
Rank of $ D $	1	2.5	2.5	4.5	4.5	6	7	8	9	10	11
Signed rank	-1	-2.5	2.5	4.5	4.5	6	7	8	9	10	11

Example: paired signed-rank test with ties (cont.)

Solution (cont.)

Hence $W^- = 1 + 2.5 = 3.5$ and $W^+ = 66 - 3.5 = 62.5$, where $n(n+1)/2 = 66$. Small values of W^- support H_1 .

Using the table. From the signed-rank table in Appendix B of the lecture notes, the one-sided 5% and 1% critical values for $n = 11$ are 13 and 7. Since $W^- = 3.5 \leq 7$, H_0 is rejected even at the 1% level. (The table is exact for untied data; with a few ties it remains a good guide.)

Example: paired signed-rank test with ties (cont.)

Solution (cont.)

Normal approximation with tie correction. $\mathbb{E}(W^-) = n(n+1)/4 = 33$.

There are two tie groups of size $t = 2$ (the 3's and the 5's), so

$$\text{Var}(W^-) = \frac{11 \cdot 12 \cdot 23}{24} - \frac{1}{48} \sum_j t_j(t_j^2 - 1) = 126.5 - \frac{2 \times 6}{48} = 126.25.$$

With continuity correction,

$$z = \frac{3.5 + 0.5 - 33}{\sqrt{126.25}} = \frac{-29}{11.236} = -2.581, \quad \mathbb{P} \approx \Phi(-2.581) = 0.0049.$$

Example: paired signed-rank test with ties (cont.)

Solution (cont.)

Exact conditional p -value. Under H_0 the $2^{11} = 2048$ sign patterns attached to the ranks $1, 2.5, 2.5, 4.5, \dots, 11$ are equally likely. The sum of the negative ranks is at most 3.5 only when the negative ranks form one of the sets $\emptyset, \{1\}, \{2.5\}, \{2.5'\}, \{1, 2.5\}, \{1, 2.5'\}$ (the two tied ranks are distinguished by a prime). Hence $\mathbb{P}(W^- \leq 3.5) = 6/2048 = 0.0029$.

Conclusion. The signed-rank test gives much stronger evidence ($p \approx 0.003$) than the sign test ($p = 0.033$) on the same data, because it uses the sizes of the differences: the two negative differences are among the smallest in absolute value. This extra power is bought with the extra assumption of symmetry.

A point estimator naturally associated with the one-sample signed-rank test, proposed by Hodges and Lehmann (1963), is the median of the $n(n+1)/2$ Walsh averages

$$\hat{\theta}_{HL} = \text{median}\left\{\frac{X_i + X_j}{2} : 1 \leq i \leq j \leq n\right\}.$$

For paired data, replace X_i by the paired differences D_i .

Confidence intervals can be obtained by inverting the signed-rank test. This illustrates how rank procedures provide estimation as well as testing.

Example 55.2 (One-sample Hodges–Lehmann estimate)

The gains in yield (quintals per hectare) on five plots after a change of fertiliser were

1.8, 2.6, 3.1, 4.0, 9.5.

Assuming a distribution symmetric about θ , estimate θ by the Hodges–Lehmann estimator and compare it with the sample mean and median.

Solution

With $n = 5$ there are $n(n + 1)/2 = 15$ Walsh averages $(X_i + X_j)/2$, $i \leq j$ (next frame).

Example: one-sample Hodges–Lehmann estimate (cont.)

Solution (cont.)

	1.8	2.6	3.1	4.0	9.5
1.8	1.80	2.20	2.45	2.90	5.65
2.6		2.60	2.85	3.30	6.05
3.1			3.10	3.55	6.30
4.0				4.00	6.75
9.5					9.50

In increasing order they are

1.80, 2.20, 2.45, 2.60, 2.85, 2.90, 3.10, **3.30**,
3.55, 4.00, 5.65, 6.05, 6.30, 6.75, 9.50.

Example: one-sample Hodges–Lehmann estimate (cont.)

Solution (cont.)

The median of the 15 Walsh averages, the 8th value, is $\hat{\theta}_{HL} = 3.30$. The sample mean is $21.0/5 = 4.20$ and the sample median is 3.10. The large value 9.5 pulls the mean upwards, but it enters only 5 of the 15 Walsh averages, so $\hat{\theta}_{HL}$ stays close to the bulk of the data while still using all the observations.

- The sign test requires only a median statement and uses n signs.
- The signed-rank test uses the order of absolute deviations and is generally more powerful under a symmetric location model.
- If symmetry is doubtful or the data have highly irregular tails, the sign test has the safer interpretation.

Practice (Lecture Notes, Chapter 3)

- Short-answer questions (Section 3.14): Q8, Q9.
- Exercises (Section 3.15): 3.19.

Next lecture

Lecture 56: Mann–Whitney–Wilcoxon test: null distribution.