ORIGIN ≡ 0

Wilcoxon Signed-Rank Test

The Signed-Rank Test is a nonparametric analog to the paired t-test utilizing more information than available in the Sign Test. The Signed-Rank test requires use of ordinal data - data that can be ordered, or ranked, according to amount of effect. However, amount or effect need not have meaning beyond order of classes of data.

Assumptions:

- **Observed values $X_{1,1}$, $X_{1,2}$, $X_{1,3}$, ... $X_{1,n}$ are a random sample exactly matched with Observed values $X_{2,1}$, $X_{2,2}$, $X_{2,3}$, ... $X_{2,n}$ across individuals 1,2,3, ... ,n.**
- **Let the value $d_i = X_{1,i} - X_{2,i}$ for each individual i be assessed as $|d_i|$ = rank order of single observations or discrete classes of observations with observed frequency.**
- **The d_i's are independent.**
- **The underlying distribution of the d_i's is continuous & symmetric but not necessarily a Normal Distribution.**
- **All d_i's have the same median**

Dermatology Example >
Rosner 2006 Ex 9.12, p. 370

	Before	After	d
1	1	9	-8
2	1	8	-7
3	1	8	-7
4	1	8	-7
5	1	7	-6
6	1	7	-6
7	1	6	-5
8	1	6	-5
9	1	5	-4
10	2	5	-3
11	2	5	-3
12	3	6	-3
13	1	4	-3
14	1	4	-3
15	3	5	-2
16	2	4	-2
17	1	3	-2
18	1	3	-2
19	1	2	-1
20	2	3	-1
21	3	4	-1
22	4	5	-1
23	5	4	1
24	6	5	1
25	4	3	1
26	2	1	1
27	2	1	1
28	3	2	1
29	5	4	1
30	6	5	1
31	2	1	1
32	2	1	1
33	3	1	2
34	4	2	2
35	5	3	2
36	6	4	2
37	5	3	2
38	4	2	2
39	4	1	3
40	5	2	3

Hypotheses:

H_0: Δ = 0 < **No population ordinal difference in median**

H_1: $\Delta \neq 0$ < **Two sided test**

Rank Data and Sum:

- **Ignore all d_i's = 0 - Don't include them in the rankings.**
- **The $|d_i|$'s are ranked (R_i = rank($|d_i|$) according to their absolute value with smallest $|d_i|$ = 1 and largest $|d_i|$ = n.**
- **Give all d_i's with same absolute value the same average rank.**
- **Count number of ties (t_j) for each group (g) of ties for the d_i's**
- **Compute the Rank Sum [RS_{pos}] of positive d_i's.**

$$d_{neg} := \begin{pmatrix} -8 \\ -7 \\ -6 \\ -5 \\ -4 \\ -3 \\ -2 \\ -1 \end{pmatrix} \quad count_{neg} := \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \\ 1 \\ 5 \\ 4 \\ 4 \end{pmatrix} \quad d_{pos} := \begin{pmatrix} 8 \\ 7 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} \quad count_{pos} := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 6 \\ 10 \end{pmatrix}$$

$R := d_{pos}$

$count := count_{neg} + count_{pos}$

$$R = \begin{pmatrix} 8 \\ 7 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} \quad count = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \\ 1 \\ 7 \\ 10 \\ 14 \end{pmatrix}$$

$i := 0 .. \text{length}(R) - 1$

$n := \sum_i count_i \qquad n = 40$

< **count in absolute value of rank (sums the positive and negative counts above)**

Averaging Absolute Value of Rank:

$$R = \begin{pmatrix} 8 \\ 7 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} \quad count = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \\ 1 \\ 7 \\ 10 \\ 14 \end{pmatrix} \quad AR := \begin{pmatrix} 40 \\ 38 \\ 35.5 \\ 33.5 \\ 32 \\ 28 \\ 19.5 \\ 7.5 \end{pmatrix} \begin{pmatrix} \text{"< average rank = sum(40)/1"} \\ \text{"< average rank = sum(37-39)/3"} \\ \text{"< average rank = sum(35-36)/2"} \\ \text{"< average rank = sum(33-34)/2"} \\ \text{"< average rank = sum(32)/2 "} \\ \text{"< average rank = sum(25-31)/7"} \\ \text{"< average rank = sum(15-24)/10"} \\ \text{"< average rank = sum(1-14)/14"} \end{pmatrix} \quad count_{pos} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 6 \\ 10 \end{pmatrix}$$

<
< multiplied
< by AR

Computing Rank Sum Positive:

$RS_{pos} := 10 \cdot 7.5 + 6 \cdot 19.5 + 2 \cdot 28$ $\qquad RS_{pos} = 248$ **< sum of the ranks for positive d_i's**

Computing Criterion for Test Statistic T:

$\frac{n \cdot (n+1)}{4} = 410$ **< criterion for test statistic T** **< Critical Value** $\qquad n = 40$

Criterion for Normal Approximation:

- IF number of non-zero d_i < 16 THEN use Special Tables e.g., Rosner Table 11 in Appendix OTHERWISE Normal Approximation may be used

$n = 40$ **< qualifies for Normal Approximation**

$t := count$

$St := \sum (t^3 - t)$ $\qquad St = 4092$

^ note use of vector sum function here that adds all elements of a vector together...

$$t = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \\ 1 \\ 7 \\ 10 \\ 14 \end{pmatrix} \quad t^3 = \begin{pmatrix} 1 \\ 27 \\ 8 \\ 8 \\ 1 \\ 343 \\ 1000 \\ 2744 \end{pmatrix} \quad t^3 - t = \begin{pmatrix} 0 \\ 24 \\ 6 \\ 6 \\ 0 \\ 336 \\ 990 \\ 2730 \end{pmatrix}$$

^ sum these for St

Test Statistic:

IF $RS_{pos} \neq \frac{n(n+1)}{4}$ **AND there ARE NO ties THEN:**

$RS_{pos} = 248$

$$T := \frac{\left[\left|RS_{pos} - \frac{n \cdot (n+1)}{4}\right| - \frac{1}{2}\right]}{\sqrt{\frac{n \cdot (n+1) \cdot (2 \cdot n + 1)}{24}}} \qquad T = 2.1708$$

IF $RS_{pos} \neq \frac{n(n+1)}{4}$ **AND there ARE ties THEN:**

$$T := \frac{\left[\left|RS_{pos} - \frac{n \cdot (n+1)}{4}\right| - \frac{1}{2}\right]}{\sqrt{\frac{n \cdot (n+1) \cdot (2 \cdot n + 1)}{24} - \frac{\sum (t^3 - t)}{48}}} \qquad T = 2.1877$$

< this one applies here, since there are ties

IF RS_{pos} = n(n+1)/4 THEN:

$T := 0$

GENERAL ALTERNATIVE TO THE ABOVE:

$j := 0 .. \text{length}(AR) - 1$

$$T := \frac{\left[\left|RS_{pos} - \frac{n \cdot (n+1)}{4}\right| - \frac{1}{2}\right]}{\sqrt{\sum_j \frac{\left[t_j \cdot \left(AR_j\right)^2\right]}{4}}} \qquad T = 2.1877$$

< this one also applies

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$C := \text{qnorm}\left(1 - \frac{\alpha}{2}, 0, 1\right) \qquad C = 1.96$

Decision Rule:

IF T > C THEN REJECT H_0 OTHERWISE ACCEPT H_0

$T = 2.1877 \qquad C = 1.96$

Probability Value:

$P := 2 \cdot (1 - \text{pnorm}(T, 0, 1)) \qquad P = 0.02869$

Prototype in R:

```
#WILCOXON SIGNED-RANK TEST
#PAIRED t-TEST ANALOG

ROS=read.table("c:/DATA/Biostatistics/RosnerEX9.12.txt")
ROS
attach(ROS)
wilcox.test(Before,After,alternative="two.sided",paired=T)
```

```
        Wilcoxon signed rank test with continuity correction

data:  Before and After
V = 248, p-value = 0.02869
alternative hypothesis: true location shift is not equal to 0

Warning message:
In wilcox.test.default(Before, After, alternative = "two.sided",  :
  cannot compute exact p-value with ties
```

^ Note: V in R's report = RS_{pos} above. P values match.