Sign and signed rank test

STAT 401 - Statistical Methods for Research Workers

Jarad Niemi

Iowa State University

7 September 2013

Rusty leaves data

	year1	year2	diff	>0
1	38	32	6	1
2	10	16	-6	0
3	84	57	27	1
4	36	28	8	1
5	50	55	-5	0
6	35	12	23	1
7	73	61	12	1
8	48	29	19	1

Rusty leaves data

	year1	year2	diff	>0
1	38	32	6	1
2	10	16	-6	0
3	84	57	27	1
4	36	28	8	1
5	50	55	-5	0
6	35	12	23	1
7	73	61	12	1
8	48	29	19	1

If the null hypothesis is true, then the final column should have 50/50 change of being 0 or 1.

Rusty leaves data

	year1	year2	diff	>0
1	38	32	6	1
2	10	16	-6	0
3	84	57	27	1
4	36	28	8	1
5	50	55	-5	0
6	35	12	23	1
7	73	61	12	1
8	48	29	19	1

If the null hypothesis is true, then the final column should have 50/50 change of being 0 or 1. In this example, we have 6 ones out of 8.

Sign test

The sign test calculates the probability of observing this many ones (or more extreme) if the null hypothesis is true.

Sign test

The sign test calculates the probability of observing this many ones (or more extreme) if the null hypothesis is true. For our one-sided hypothesis, that is the probability of observing 6, 7, or 8 ones.

Sign test

The sign test calculates the probability of observing this many ones (or more extreme) if the null hypothesis is true. For our one-sided hypothesis, that is the probability of observing 6, 7, or 8 ones. This is

$$\binom{8}{6}0.5^8 + \binom{8}{7}0.5^8 + \binom{8}{8}0.5^8 = 0.14$$

Sign test

The sign test calculates the probability of observing this many ones (or more extreme) if the null hypothesis is true. For our one-sided hypothesis, that is the probability of observing 6, 7, or 8 ones. This is

$$\binom{8}{6}0.5^8 + \binom{8}{7}0.5^8 + \binom{8}{8}0.5^8 = 0.14$$

Although conceptually simple, this test is not very powerful and therefore is used rarely.

Signed rank test

Also known as the Wilcoxon signed rank test:

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.
- Order the absolute differences from smallest to largest and assign them their ranks.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.
- Order the absolute differences from smallest to largest and assign them their ranks.
- Calculate S: the sum of the ranks from the pairs for which the difference is positive.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.
- Order the absolute differences from smallest to largest and assign them their ranks.
- Calculate S: the sum of the ranks from the pairs for which the difference is positive.
- Calculate $E[S] = n(n+1)/4$ where n is the number of pairs.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.
- Order the absolute differences from smallest to largest and assign them their ranks.
- Calculate S: the sum of the ranks from the pairs for which the difference is positive.
- Calculate $E[S] = n(n+1)/4$ where n is the number of pairs.
- Calculate $SD[S] = [n(n+1)(2n+1)/24]^{1/2}$.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.
- Order the absolute differences from smallest to largest and assign them their ranks.
- Calculate S: the sum of the ranks from the pairs for which the difference is positive.
- Calculate $E[S] = n(n+1)/4$ where n is the number of pairs.
- Calculate $SD[S] = [n(n+1)(2n+1)/24]^{1/2}$.
- Calculate $(S + c - E[S])/SD[S]$ where c is the appropriate continuity correction.

Signed rank test

Also known as the Wilcoxon signed rank test:

- Compute the difference in each pair.
- Drop zeros from the list.
- Order the absolute differences from smallest to largest and assign them their ranks.
- Calculate S: the sum of the ranks from the pairs for which the difference is positive.
- Calculate $E[S] = n(n+1)/4$ where n is the number of pairs.
- Calculate $SD[S] = [n(n+1)(2n+1)/24]^{1/2}$.
- Calculate $(S + c - E[S])/SD[S]$ where c is the appropriate continuity correction.
- Calculate the pvalue from a standard normal table.

Signed rank test

	year1	year2	diff	>0	absdiff	rank
5	50	55	-5	0	5	1
1	38	32	6	1	6	2.5
2	10	16	-6	0	6	2.5
4	36	28	8	1	8	4
7	73	61	12	1	12	5
8	48	29	19	1	19	6
6	35	12	23	1	23	7
3	84	57	27	1	27	8

Signed rank test

	year1	year2	diff	>0	absdiff	rank
5	50	55	-5	0	5	1
1	38	32	6	1	6	2.5
2	10	16	-6	0	6	2.5
4	36	28	8	1	8	4
7	73	61	12	1	12	5
8	48	29	19	1	19	6
6	35	12	23	1	23	7
3	84	57	27	1	27	8

- $S = 32.5$

Signed rank test

	year1	year2	diff	>0	absdiff	rank
5	50	55	-5	0	5	1
1	38	32	6	1	6	2.5
2	10	16	-6	0	6	2.5
4	36	28	8	1	8	4
7	73	61	12	1	12	5
8	48	29	19	1	19	6
6	35	12	23	1	23	7
3	84	57	27	1	27	8

- $S = 32.5$
- $E[S] = 18$

Signed rank test

	year1	year2	diff	>0	absdiff	rank
5	50	55	-5	0	5	1
1	38	32	6	1	6	2.5
2	10	16	-6	0	6	2.5
4	36	28	8	1	8	4
7	73	61	12	1	12	5
8	48	29	19	1	19	6
6	35	12	23	1	23	7
3	84	57	27	1	27	8

- $S = 32.5$
- $E[S] = 18$
- $SD[S] = 7.14$

Signed rank test

	year1	year2	diff	>0	absdiff	rank
5	50	55	-5	0	5	1
1	38	32	6	1	6	2.5
2	10	16	-6	0	6	2.5
4	36	28	8	1	8	4
7	73	61	12	1	12	5
8	48	29	19	1	19	6
6	35	12	23	1	23	7
3	84	57	27	1	27	8

- $S = 32.5$
- $E[S] = 18$
- $SD[S] = 7.14$
- $Z = 1.96$ (with continuity correction of -0.5)

Signed rank test

	year1	year2	diff	>0	absdiff	rank
5	50	55	-5	0	5	1
1	38	32	6	1	6	2.5
2	10	16	-6	0	6	2.5
4	36	28	8	1	8	4
7	73	61	12	1	12	5
8	48	29	19	1	19	6
6	35	12	23	1	23	7
3	84	57	27	1	27	8

- $S = 32.5$
- $E[S] = 18$
- $SD[S] = 7.14$
- $Z = 1.96$ (with continuity correction of -0.5)
- $p = 0.02$

SAS code for paired nonparametric test

```
DATA leaves;
  INPUT tree year1 year2;
  diff = year1-year2;
  DATALINES;
1 38 32
2 10 16
3 84 57
4 36 28
5 50 55
6 35 12
7 73 61
8 48 29
;

PROC UNIVARIATE DATA=leaves;
    VAR diff;
    RUN;
```

SAS code for paired nonparametric tests

The UNIVARIATE Procedure
Variable: diff

Moments

N	8	Sum Weights	8
Mean	10.5	Sum Observations	84
Std Deviation	12.2007026	Variance	148.857143
Skewness	-0.1321468	Kurtosis	-1.2476273
Uncorrected SS	1924	Corrected SS	1042
Coeff Variation	116.197167	Std Error Mean	4.31359976

Basic Statistical Measures

Location		Variability	
Mean	10.50000	Std Deviation	12.20070
Median	10.00000	Variance	148.85714
Mode	.	Range	33.00000
		Interquartile Range	20.50000

Tests for Location: Mu0=0

Test	-Statistic-		-----p Value------	
Student's t	t	2.434162	Pr > |t|	0.0451
Sign	M	2	Pr >= |M|	0.2891
Signed Rank	S	14.5	Pr >= |S|	0.0469

Conclusion

Removal of red cedar trees within 100 yards is associated with a significant reduction in rusty apple leaves (Wilcoxon signed rank test, p=0.023).