The single-sample Wilcoxon signed-rank test

Starter

1. **(Review of last lesson)**

   Carry out single-sample sign tests, at the 5% level, on the following data sets against

   H_0 : Median is 30

   H_1 : Median is more than 30.

   State your conclusions clearly.

   (a) 26 28 29 32 33 35 36 39 40 41 43 44

   (b) 21 23 24 37 38 40 41 44 45 46 48 49

Notes

The starter demonstrates the limitations of the sign test — the data in (b) is a lot further away from the median but there is no chance of a different conclusion to (a) since the same number of values are above and below the median.

The ***single-sample Wilcoxon signed-rank test*** is a ***more sophisticated*** test than the ***single-sample sign test*** because it takes into account the ***size of the deviations*** from the medians, as well as the signs i.e. values further from the median make a greater contribution than values close to the median.

Required assumption for single-sample Wilcoxon signed-rank test

A further difference from the single-sample sign test is that an assumption must be made about the distribution. The ***population must be symmetrical*** so that the mean and median are equal

H_0 : population median mark is __

H_1 : population median mark is __

If the test statistic is ***less than or equal to*** the critical value, ***reject*** H_0.

If the test statistic is ***more than*** the critical value, ***do not reject*** H_0.

Success Criteria

1. Cross out values that equal the median.
2. Calculate "difference = data value — median" for each data value.
3. Highlight the negative values.
4. Ignoring the signs, rank the differences — smallest = rank 1.
5. With tied ranks, take the average of the ranks.
6. Make these numbers positive.
7. Rank the differences.
8. Sum the positive ranks, W_+.
9. Sum the negative ranks, W_-.
10. Identify T, the smaller of W_+ and W_-.
11. Compare T with the critical value in the table.

In the event of ***tied ranks***, the ***average rank*** is assigned.

A table like the one below, with as many columns as data values, can be useful:

Value			
Difference			
|Difference|			
Rank			
Signed rank			

Knowing the sum of the ranks

The formula $\frac{1}{2}n(n+1)$ gives the sum of the first n natural numbers i.e. $1+2+3+4+\ldots+n$.

This can be used to check your values of W_+ and W_- since $W_+ + W_- = \frac{1}{2}n(n+1)$.

E.g. 1 State the critical values for the following single-sample Wilcoxon signed-ranks tests:
(a) a one-tailed test at the 1% level with 9 values
(b) a one-tailed test at the 5% level with 14 values
(c) a two-tailed test at the 5% level with 16 values

Working: (a) 3

E.g. 2 A random sample of times, in hours, for cars driving from Perth to Melbourne is given.

$$63.6 \quad 55.2 \quad 54.1 \quad 68.1 \quad 52.3 \quad 47.1 \quad 125.9 \quad 60.8 \quad 61.4 \quad 53.6$$

Assuming that the distribution of times is symmetrical, carry out a single-sample Wilcoxon signed-rank test to test the null hypothesis that the median time for this journey is 55 hours against the alternative hypothesis that the median time for this journey is greater than 55 hours. Use a 5% significance level.

Working: H_0 : the population median is 55 hours
H_1 : the population median is greater than 55 hours

Value	63.6	55.2	54.1	68.1	52.3	47.1	125.9	60.8	61.4	53.6
Difference	8.6	0.2	−0.9	13.1	−2.7	−7.9	70.9	5.8	6.4	−1.4
|Difference|	8.6	0.2	0.9	13.1	2.7	7.9	70.9	5.8	6.4	1.4
Rank	8	1	2	9	4	7	10	5	6	3
Signed rank	8	1	−2	9	−4	−7	10	5	6	−3

$W_+ = 39$ and $W_- = 16$

Check: when $n = 10$, $\frac{1}{2} \times 10 \times (10+1) = 55 = 39 + 16$ ✓

$T = 16$ (smallest value)

From tables, the critical value for a one-tail test at the 5% level with 10 values is 10.

Since $T = 16 \nleq 10 = CV$, we do not reject H_0.

There is evidence to suggest that the population median is 55 hours.

E.g. 3 A bus is timetabled to arrive at $15\ 35$. Its actual arrival time is noted on 10 randomly chosen occasions and the numbers of minutes late or early are recorded, correct to the nearest half-minute. The data is recorded below with positive entries indicating a late time and a negative entry showing an early time.

$$0.5 \quad 3.0 \quad 2.5 \quad -1.0 \quad -3.5 \quad 4.0 \quad 10.0 \quad 5.0 \quad -2.0 \quad 7.5$$

Assuming that the distribution is symmetrical, carry out a single-sample Wilcoxon signed-rank test at the 2.5% significance level to test the claim that the bus is late more often than not.

E.g. 4 Becotide inhalers for asthmatics should deliver 50 mg of the active ingredient per puff. In a test, 14 puffs from randomly selected inhalers were tested and the amount of active ingredient delivered was:

43.1	47.3	52.4	51.2	44.7	50.8
51.8	46.7	52.0	50.5	47.7	45.3

(a) Explain why a Wilcoxon test is preferable to a sign test in this context.
(b) Use a single-sample signed-rank test at the 10% level to test whether the inhalers are delivering the correct amount of active ingredient.

Video (password needed): Single-sample Wlicoxon signed-rank test
Video: Single-sample Wlicoxon signed-rank test
Video: Wilcoxon signed-rank test

Solutions to Starter and E.g.s

Exercise
p51 4B Qu 1i, 2-4, (5 red)

Summary
The ***population must be symmetrical*** so that the mean and median are equal
H_0 : population median mark is __
H_1 : population median mark is __

If the test statistic is ***less than or equal to*** the critical value, ***reject*** H_0.
If the test statistic is ***more than*** the critical value, ***do not reject*** H_0.

Success Criteria
1. Cross out values that equal the median.
2. Calculate "difference = data value — median" for each data value.
3. Highlight the negative values.
4. Ignoring the signs, rank the differences — smallest = rank 1.
5. With tied ranks, take the average of the ranks
6. Make these numbers positive.
7. Rank the differences.
8. Sum the positive ranks, W_+.
9. Sum the negative ranks, W_-.
10. Identify T, the smaller of W_+ and W_-.
11. Compare T with the critical value in the table.

In the event of ***tied ranks***, the ***average rank*** is assigned.

A table like the one below, with as many columns as data values, can be useful:

Value			
Difference			
|Difference|			
Rank			
Signed rank			

Check: $W_+ + W_- = \frac{1}{2}n(n+1)$