Matched-pairs (or paired-sample) tests

Starter

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

A machine is designed to produce rods of median length 2 cm. After the machine has been installed the first nine rods are measured and their lengths are the following:

$$1.89 \quad 1.92 \quad 2.05 \quad 1.88 \quad 1.96 \quad 1.97 \quad 2.01 \quad 1.94 \quad 1.90$$

Test, at the $10\,\%$ significance level, whether the median is different to 2 cm:

(a) using a single-sample sign test

(b) using a single-sample Wilcoxon signed-rank test.

Working: (a)

H_0 : Median $= 2$ cm

H_1 : Median $\neq 2$ cm

Let X be the number of values different to 2 so $X \sim$ B$(9, 0.5)$

Signs of deviations: $- \quad - \quad + \quad - \quad - \quad - \quad + \quad - \quad -$

$p = P(X \geq 7) = P(X \leq 2) = 0.0898$

Since $p = 0.0898 \nleq 0.05$, we do not reject H_0.

There is evidence to suggest that the median is 2 cm.

(b)

H_0 : Median $= 2$ cm

H_1 : Median $\neq 2$ cm

Value	1.89	1.92	2.05	1.88	1.96	1.97	2.01	1.94	1.90
Difference	−0.11	−0.88	0.05	−0.12	−0.04	−0.03	0.01	−0.06	−0.10
|Difference|	0.11	0.88	0.05	0.12	0.04	0.03	0.01	0.06	0.10
Rank	8	6	4	9	3	2	1	5	7
Signed rank	−8	−6	4	−9	−3	−2	1	−5	−7

$W_+ = 5$ and $W_- = 40$

Check: when $n = 9$, $\frac{1}{2} \times 9 \times (9+1) = 45 = 5 + 40$ ✓

$T = 5$ (smallest value)

From tables, the two-tail critical value at the $10\,\%$ level is 8.

Since $T = 5 \leq 8 = CV$, we reject H_0.

There is evidence to suggest that the median is not 2 cm.

E.g. 1 The table below shows the times taken by a random sample of people to perform a simple task on their first and second attempts. Test, at the $10\,\%$ significance level, whether most people take less time on the second attempt than on the first attempt.

Person	A	B	C	D	E	F	G	H
1st attempt	6.3	3.5	7.1	3.7	8.4	3.9	4.7	5.2
2nd attempt	5.1	3.4	6.2	4.5	7.3	4.0	3.6	5.1

Working:

H_0 : the time taken on the 2nd attempt is the same as on the 1st attempt
H_1 : the time taken on the 2nd attempt is lower than on the 1st attempt
By H_0, X, the number of "—" signs is distributed such that $X \sim \mathrm{B}(8, 0.5)$
The signs of the differences between the 1st and 2nd attempts are.
$- \quad - \quad - \quad + \quad - \quad + \quad - \quad -$
$p = P(X \geq 6) = P(X \leq 2) = 0.145$
Since $p = 0.145 > 0.10$, we do not reject H_0.
There is no evidence to suggest that there is an improvement in times from the first attempt to the second attempt.

E.g. 2 The numbers of male and female residents in eleven randomly selected villages are shown in the table below

Village	A	B	C	D	E	F	G	H	I	J	K
Male	196	169	335	220	298	215	461	250	370	355	382
Female	220	171	361	248	300	237	434	325	451	345	401

Use a binomial sign test to determine the validity of the hypothesis that there are more female than male residents in villages. Make clear your hypotheses and conduct the test at the $5\,\%$ significance level.

Working:

H_0 : the number of male and female residents in villages is equal
H_1 : most villages have more female than male residents
By H_0, X, the number of "—" signs is distributed such that $X \sim \mathrm{B}(8, 0.5)$
The signs of the differences between the 1st and 2nd attempts are.
$+ \quad + \quad + \quad + \quad + \quad + \quad - \quad + \quad + \quad - \quad +$
$p = P(X \geq 9) = P(X \leq 2) = 0.0327$
Since $p = 0.0327 < 0.05$, we reject H_0.
There is no evidence to suggest that most villages have more female than male residents.

E.g. 3 Ten people enrolled on a new slimming course. Their weights in kilograms before and after the course are shown in the table below.

Person	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Before	75.4	78.1	79.7	70.3	72.0	74.1	78.5	74.9	70.3	72.9
After	70.9	71.3	69.5	73.2	72.1	72.0	71.6	73.1	70.8	71.6

Test at the 5% level whether the course is effective.

Working: H_0 : the course was not effective in helping people to lose weight
H_1 : the course was effective in helping people to lose weight

Person	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Before	75.4	78.1	79.7	70.3	72.0	74.1	78.5	74.9	70.3	72.9
After	70.9	71.3	69.5	73.2	72.1	72.0	71.6	73.1	70.8	71.6
Differences	4.5	6.8	10.2	−2.9	−0.1	2.1	6.9	1.8	−0.5	1.3
|Differences|	4.5	6.8	10.2	2.9	0.1	2.1	6.9	1.8	0.5	1.3
Rank	7	8	10	6	1	5	9	4	2	3
Signed rank	7	8	10	−6	−1	5	9	4	−2	3

$W_+ = 46$ and $W_- = 9$

Check: when $n = 10$, $\frac{1}{2} \times 10 \times (10+1) = 55 = 46 + 9$ ✓

$T = 9$ (smallest value)

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

Since $T = 9 \leq 10 = CV$, we reject H_0.

There is evidence to suggest that the course was effective in helping people to lose weight.

E.g. 4 A team of scientists believe they have found a drug that improves memory in older people. They test this on a group of nine pairs of twins by asking a long series of questions about their childhood. One twin had taken the drug and the other had not. The number of correct answers is below:

Pair of twins	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Twin taking drug	94	138	66	142	137	90	123	154	141
Twin not taking drug	83	121	75	157	118	92	105	134	127

Carry out an appropriate Wilcoxon test, at the 5% significance level, to determine whether the drug has improved the ability to recall information. State any assumption that is necessary to justify the use of the test.

Working: H_0 : the drug does not improve memory
H_1 : the drug improves memory

Pair of twins	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Twin taking drug	94	138	66	142	137	90	123	154	141
Twin not taking drug	83	121	75	157	118	92	105	134	127
Differences	11	17	−9	−15	19	−2	18	20	14
|Differences|	11	17	9	15	19	2	18	20	14
Rank	3	6	2	5	8	1	7	9	4
Signed rank	3	6	−2	−5	8	−1	7	9	4

$W_+ = 37$ and $W_- = 8$
Check: when $n = 9$, $\frac{1}{2} \times 9 \times (9+1) = 45 = 37 + 8$ ✓
$T = 8$ (smallest value)
From tables, the critical value for a one-tail test at the 5% level with 9 values is 8.
Since $T = 8 \leq 8 = CV$, we reject H_0.
There is evidence to suggest that the drug improves memory.
The assumption is that the distribution of the differences is symmetrical.

E.g. 5 As part of a paired-sample Wilcoxon signed rank test, the differences in the paired data were given the signed ranks below. The $(+)$ or $(-)$ beside each rank indicates whether the difference was positive or negative.

$1(-)$ $2(-)$ $3(+)$ $4(-)$ $5(+)$ $6(+)$ $7(+)$ $8(+)$ $9(+)$ $10(+)$

Test whether there is a difference between the population medians, using a two-tailed test at the 5% significance level.

Working:

H_0 : there is no difference between the population medians
H_1 : there is a difference between the population medians
$W_+ = 48$ and $W_- = 7$
Check: when $n = 10$, $\frac{1}{2} \times 10 \times (10 + 1) = 55 = 48 + 7$ ✓
$T = 7$ (smallest value)
From tables, the critical value for a one-tail test at the 5% level with 10 values is 8.
Since $T = 7 \leq 8 = CV$, we reject H_0.
There is evidence to suggest that there is a difference between the population medians.

Video (password needed): Paired and unpaired samples
Video (password needed): Wilcoxon matched-pairs signed-test
Video (Paired): Worked example of Wilcoxon signed-rank test
Video: Wilcoxon signed-rank test (paired data)
Video: Paired Wlicoxon signed-rank test

Solutions to Starter and E.g.s

Exercise
p54 4C Qu 1-5, (6, 7 red)