

# CIMT Further Statistics - p84 - Ex4F

1.

Region	A	B	C	D	E	F	G	H	I	J	K
Before	2.4	2.6	3.9	2.0	3.2	2.2	3.3	2.1	3.1	2.2	2.8
After	3.0	2.5	4.0	4.1	4.8	2.0	3.4	4.0	3.3	4.2	3.9
After-before	0.6	-0.1	0.1	2.1	1.6	-0.2	0.1	1.9	0.2	2	1.1
Rank	6	2	2	11	8	4.5	2	9	4.5	10	7

Assumptions: distribution of differences in sales revenue is symmetrical

Justification: paired data

$H_0$ : population median difference,  $\eta_D = 0$  (campaign had no effect)

$H_1$ : population median difference,  $\eta_D > 0$  (campaign increased sales)

Assume  $H_0$  to be true.

where difference = after - before.

$\alpha = 5\%$

1-tail test

From table above  $W_- = 2 + 4.5 = 6.5$

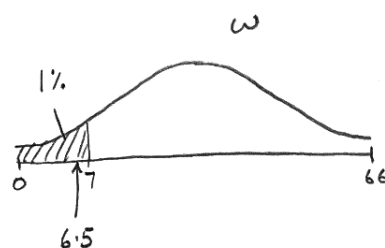
$$W_+ = 6 + 2 + 11 + 8 + 2 + 9 + 4.5 + 10 + 7 = 59.5 \quad \left. \vphantom{W_+} \right\} 66 = \frac{1}{2} \times 11 \times 12 \quad \checkmark \text{check.}$$

and  $n = 11$

now  $W = \min(W_-, W_+)$   
 $= 6.5$

we are interested in  $P(W \leq 6.5)$

From tables, for  $n = 11$ ,  
 $P(W \leq 13) = 0.05$   
 $P(W \leq 10) = 0.025$   
 $P(W \leq 7) = 0.01$   
 $P(W \leq 5) = 0.005$



so  $P(W \leq 6.5)$  is between 0.005 and 0.01

we are in the 5% critical region (we're also in the 1% critical region)

so, we have evidence to reject  $H_0$  at the 5% level and conclude that the population median difference is greater than zero. This means that the advertising campaign seems to have increased sales.

2. Assumptions: distribution of differences in clay pigeon shooting scores is symmetrical

Justification: paired data

$H_0$ : population median difference,  $\eta_D = 0$  (guns equally good)

$H_1$ : population median difference,  $\eta_D > 0$  (crackshot more accurate)

Assume  $H_0$  to be true.

where difference = crackshot - fastfire.

$\alpha = 5\%$ . 1 tail test

Competitor	A	B	C	D	E	F	G	H	I	J.
Crackshot	93	99	90	86	85	94	87	91	96	79
Fastfire	87	91	86	87	78	95	89	84	88	74
C-F	6	8	4	-1	7	-1	-2	7	8	5
rank	6	9.5	4	1.5	7.5	1.5	3	7.5	9.5	5

$n = 10$ .

$$W_- = 1.5 + 1.5 + 3 = 6$$

$$W_+ = 6 + 9.5 + 4 + 7.5 + 7.5 + 9.5 + 5 = 49 \quad \left. \vphantom{W_+} \right\} 6 + 49 = 55 = \frac{1}{2} \times 10 \times 11 \quad \checkmark \text{ check}$$

$$\text{now } W = \min(W_-, W_+) \\ = 6.$$

we want  $P(W \leq 6)$  for  $n = 10$

From tables

$$P(W \leq 10) = 0.5$$

$$P(W \leq 8) = 0.025$$

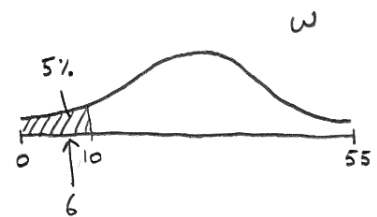
$$P(W \leq 5) = 0.01$$

So  $P(W \leq 6)$  is between 0.01 and 0.025

So at 5%, we are in critical region

Hence, we have evidence to reject  $H_0$  and conclude that the median difference is greater than zero, which suggests that Crackshot is more accurate than

Fastfire.



### Ex4F no. 3

Assumptions: distribution of differences in asthmatic index is symmetrical

Justification: paired data

$H_0$ : population median difference,  $\mu_D = 0$  (drug has no effect)

$H_1$ : population median difference,  $\mu_D > 0$  (drug reduces asthmatic index)

Assume  $H_0$  to be true

where difference = placebo - drug

$\alpha = 5\%$

1-tail test

Subject	1	2	3	4	5	6	7	8	9	10	11	12
Drug	28	31	17	18	31	12	33	24	18	25	19	17
Placebo	32	33	23	26	34	17	30	24	19	23	21	24
P-D	4	2	6	8	3	5	-3	0	1	-2	2	7
rank	7	3	9	11	5.5	8	5.5		1	3	3	10

$n = 11$  (as discount the zero)

so  $W_- = 5.5 + 3 = 8.5$

$W_+ = 7 + 3 + 9 + 11 + 5.5 + 8 + 1 + 3 + 10 = 57.5$  } sum to 66 =  $\frac{1}{2} \times 11 \times 12$  ✓ check

so  $W = \min(W_-, W_+)$

$= 8.5$

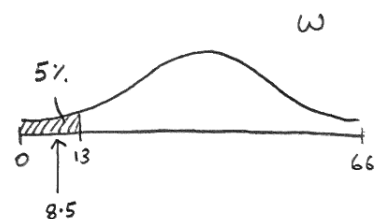
we want to know  $P(W \leq 8.5)$

from tables for  $n=11$ , we have

$$P(W \leq 13) = 0.05$$

$$P(W \leq 10) = 0.025$$

$$P(W \leq 7) = 0.01$$



so  $P(W \leq 8.5)$  is between 0.025 and 0.01

so we are inside 5% critical region

Hence we have evidence to reject  $H_0$  and conclude that the median difference is greater than zero, which suggests that the drug appears to reduce the asthmatic index.

Assumptions : distribution of differences in number of uses of the word "nice" is symmetrical.

Justification: paired data

$H_0$ : population median differences,  $\eta_D = 0$  (girls can accurately remember)

$H_1$ : population median differences,  $\eta_D \neq 0$  (girls cannot accurately remember)

2-tail test

where difference = true - recorded.

$\alpha = 5\%$

Assume  $H_0$  to be true

Girl	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
True	12	20	1	8	0	12	12	17	6	5	24	23	10	18	16
Recorded	9	19	3	14	4	12	16	14	5	9	20	16	11	17	19
T-R	3	1	-2	-6	-4	0	-4	3	1	-4	4	7	-1	1	-3
rank	7	2.5	5	13	10.5		10.5	7	2.5	10.5	10.5	14	2.5	2.5	7

so  $n = 14$

$$W_- = 5 + 13 + 10.5 + 10.5 + 10.5 + 2.5 + 7 = 59$$

$$W_+ = 7 + 2.5 + 7 + 2.5 + 10.5 + 14 + 2.5 = 46$$

$$\left. \begin{array}{l} W_- = 59 \\ W_+ = 46 \end{array} \right\} 105 = \frac{1}{2} \times 14 \times 15 \quad \checkmark \text{ check}$$

$$\begin{aligned} \text{so } W &= \min(W_-, W_+) \\ &= 46 \end{aligned}$$

we are interested in  $P(W \leq 46)$  and a 2 tail test

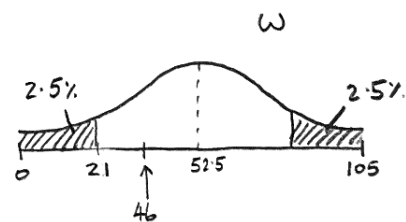
so we either compare  $P(W \leq 46)$  to 2.5%.

or compare  $2 \times P(W \leq 46)$  to 5%.

from tables, for  $n = 14$ ,

$$P(W \leq 25) = 0.05$$

$$P(W \leq 21) = 0.025 \text{ etc.}$$



so, it's clear that  $P(W \leq 46)$  is more than 0.05

and  $2 \times P(W \leq 46)$  is more than 0.10

Hence we are not in either 2.5% tail's critical region

So we do not have evidence to reject  $H_0$ . We do not have evidence that the population median difference is not zero.

This tells us that girls seem to accurately remember the frequency with which they use a particular word.

Assumptions: distribution of differences in the methods is symmetrical

Justification: paired data.

$H_0$ : population median difference,  $\eta_D = 0$  (methods give same results)

$H_1$ : population median difference,  $\eta_D \neq 0$  (methods give different results)

where difference = method 1 - method 2.

Assume  $H_0$  to be true.

2-tail test

$\alpha = 5\%$

Comparative dotplots suggests that method 1 gives slightly higher results than method 2.

Adult	1	2	3	4	5	6	7	8	9	10
Method 1	204	238	209	277	197	226	203	131	282	76
Method 2	199	230	198	253	180	209	213	137	250	82
M1-M2	5	8	11	24	17	17	-10	-6	32	-6
rank	1	4	6	9	7.5	7.5	5	2.5	10	2.5

so  $n = 10$

$$W_- = 5 + 2.5 + 2.5 = 10$$

$$W_+ = 1 + 4 + 6 + 9 + 7.5 + 7.5 + 10 = 45 \quad \left. \vphantom{W_+} \right\} 55 = \frac{1}{2} \times 10 \times 11 \quad \checkmark \text{ check}$$

$$\text{so } W = \min(W_-, W_+) = 10$$

we want  $P(W \leq 10)$  for a 2 tail test

so we either compare  $P(W \leq 10)$  to 2.5%.

or compare  $2P(W \leq 10)$  to 5%.

from tables, for  $n=10$ ,

$$P(W \leq 10) = 0.05$$

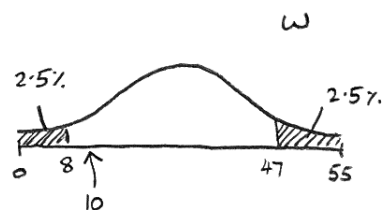
$$P(W \leq 8) = 0.025$$

so  $2P(W \leq 10) = 0.1$  which is more than 5%.

so we are not in the critical tails of the distribution

Hence we do not have evidence to reject  $H_0$ , and so we do not have evidence to suggest that population median difference is not zero.

This suggests that the two methods give the same result.



Assumptions: distribution of differences in amount of leather wear is symmetrical

Justification: paired data

$H_0$ : population median difference,  $\eta_D = 0$  (wear on boots is the same)

$H_1$ : population median difference,  $\eta_D \neq 0$  (wear on boots is different)

Assume  $H_0$  to be true.

where difference =  $A - B$ .

2-tail test

$\alpha = 5\%$

Soldier

A	5.4	2.6	4.3	1.1	3.3	6.6	4.4	3.5	1.2	1.3	4.8	1.2	2.8	2.0	6.1
B	4.7	3.2	3.8	2.3	3.6	7.2	4.4	3.9	1.9	1.2	5.8	2.0	3.7	1.8	6.1
A-B	0.7	-0.6	0.5	-1.2	-0.3	-0.6		-0.4	-0.7	0.1	-1	-0.8	-0.9	0.2	
Rank	8.5	6.5	5	13	3	6.5		4	8.5	1	12	10	11	2	

so  $n = 13$

$$W_- = 6.5 + 13 + 3 + 6.5 + 4 + 8.5 + 12 + 10 + 11 = 74.5 \quad \left. \begin{array}{l} \\ \end{array} \right\} 91 = \frac{1}{2} \times 13 \times 14 \quad \checkmark \text{ check.}$$

$$W_+ = 8.5 + 5 + 1 + 2 = 16.5$$

$$W = \min(W_-, W_+) = 16.5$$

we want  $P(W \leq 16.5)$  for a 2-tail test

so we either compare  $P(W \leq 16.5)$  to 2.5%.

or  $2P(W \leq 16.5)$  to 5%.

from tables, for  $n=13$ ,

$$P(W \leq 21) = 0.05$$

$$P(W \leq 17) = 0.025$$

$$P(W \leq 12) = 0.01$$

so  $P(W \leq 16.5)$  is between 0.025 and 0.01

so  $2P(W \leq 16.5)$  is between 0.05 and 0.02

Hence as  $2P(W \leq 16.5)$  is less than 0.05, we are in critical region

so we have evidence to reject  $H_0$  and conclude that the median difference is not zero, which means that the wear on the boots is different.

{ Furthermore, from inspection of the <sup>boxplots</sup> dotplots, it would appear that Leather A }  
is harder and wears down less. than Leather B

