Mann-Whitney Test Practice Exercise

Two examples and four practice questions (small sample sizes)

New techniques and methods to note are highlighted in yellow.

Example 1

There are 10 patients, five of whom are treated for their terminal illness and five of whom are not. The figures give how long they lived for in years, from the time of their contraction of the disease.

Treated (T)	4.2	6.5	7.9	13.2	17.8
Control (C)	0.4	1.2	2.9	5.6	6.7

Use the Mann-Whitney test to determine whether Treatment has a beneficial effective.

Solution to Example 1

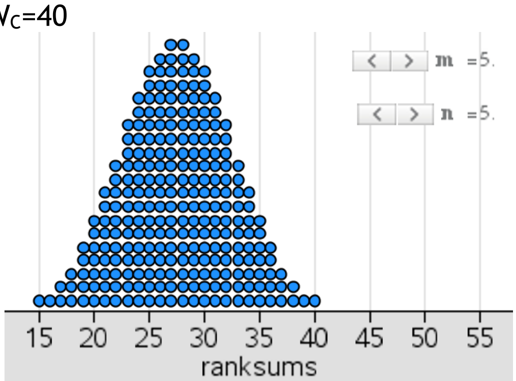
Justification for using Mann-Whitney: non-paired data
Assumptions: distributions of lifetimes have the same shape and spread

H_0: population $median_T$ = population $median_C$ (treatment has no effect)
H_1: population $median_T$ > population $median_C$ (treatment has a positive effect)

Assume H_0 to be true

We would reject H_0 in favour of H_1 for low values of W_C, or high values of W_T
After ranking $W_T=37$, $W_C=18$.
Size of data groups: m=n=5, so we can pick either rank sum to use.
We choose Wc as it has the lower value

minimum $W_C=15$, maximum $W_C=40$



We are interested in $P(W_C \le 18)$
From tables, for n=m=5, 5% critical value is 19
So we are in the critical region (as $18 \le 19$) which is equivalent to $P(W \le 18) < 5\%$
So we reject H_0.

Hence we have evidence to suggest that the treated population median is greater than the control population median, which means that treatment appears to have a positive effect.
population medians are data sets do not have the same median, so treatment has an effect.
As $W_T > W_C$, we conjecture that the treatment has a positive effect on lifespan.

Example 2:
In 1965 Capital Punishment (aka 'the death penalty') was abolished in the UK.
Some people were concerned that the removal of such a deterrent would cause an increase in the murder rate.
The figures below are for the number of murder victims in England and Wales before 1965 and after.
Perform a Mann-Whitney test to determine whether the murder rate increased after 1965.

Pre-1965	144	135	123	118	129	122
Post-1965	148	119	136	177		

Solution to Example 2

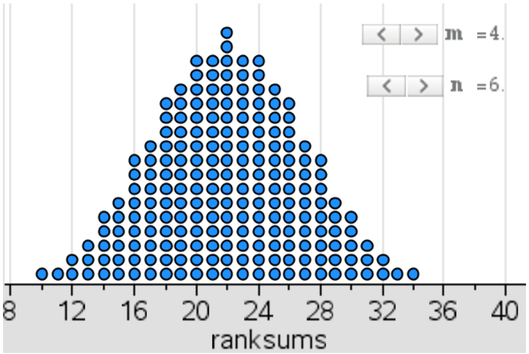
Justification for using Mann-Whitney: non-paired data
Assumptions: distributions of number of murder victims have the same shape and spread

H_0: population median$_{POST}$ = population median$_{PRE}$ (no change in murder rate)
H_1: population median$_{POST}$ > population median$_{PRE}$ (murder rate increased)

Assume H_0 to be true

We would reject H_0 in favour of H_1 for high values of W_{POST}, or small values of W_{PRE}

After ranking $W_{PRE}=27$, $W_{POST}=28$. Size of data groups: m=4, n=6
minimum $W_{POST}=10$ and maximum $W_{POST}=34$



We are interested in $P(W_{POST} \geq 28)$
By symmetry of the distribution, $P(W_{POST} \geq 28) = P(W_{POST} \leq 16)$.
From tables, for m=4, n=6, 5% critical value is 13
We are not in the critical region (as $13 \leq 16$) which is equivalent to $P(W_{POST} \leq 16) > 5\%$, so there is no evidence to reject H_0.

Hence we do not have evidence to suggest that the median population murder rates before and after 1965 did not remain equal, and thus the removal of capital punishment seems not to affect the murder rate.

The next pages contain 4 further questions, with worked solutions, for you to consolidate this process.

Question 1

1a) In Winter 1980, 6 weekdays and 4 Sundays were chosen at random. The volume of traffic for each of these days was as follows:

Weekdays	436	429	440	413	444	452
Sunday	431	424	411	392		

Using this information, test the hypothesis at the 5% level that there is less traffic on Sundays in the Winter.

b) In Summer 1980, a similar experiment gave results as follows:

Weekdays	553	546	534	517	519	545
Sunday	562	549	548	521		

Is there less traffic on Sundays in the summer?

Solution 1

1a)
Justification for using Mann-Whitney: non-paired data
Assumption: distributions of traffic data for weekdays and Sundays have the same shape and spread

H_0: population median Sunday traffic = population median weekday traffic (traffic equally busy on all days)
H_1: population median Sunday traffic < population median weekday traffic (less traffic on Sundays in winter)

Let α=5%, one-tail test.

Assume H_0 to be true

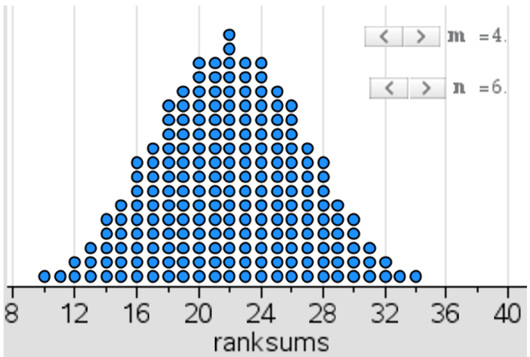
Data	392	411	413	424	429	431	436	440	444	452
Type	S	S	W	S	W	S	W	W	W	W
Rank	1	2	3	4	5	6	7	8	9	10

We would reject H_0 if we have a low value of W_s ie lots of Sundays with low traffic count, compared to weekdays.

W_s = sum of ranks of Sunday data m=4
W_w = sum of ranks of Weekday data n=6

So W_s =13 and W_w =42
Minimum rank sum of W_s = $(^1/_2)\times4\times5$=10
Maximum rank sum of W_s = 34

As m<n, we use W_s



We are interested in P($W_s \leq$ 13)
From tables, for m=4, n=6....
5% critical value is 13

Hence, at the 5% level, we would reject H_0 as W_s =13 $\leq$ 13, the 5% critical value

So we have evidence to suggest that the population median amount of Sunday traffic is less than the population median amount of Weekday traffic, which suggest that there is less traffic on Sundays in Winter.

1b)
Justification for using Mann-Whitney: non-paired data
Assumption: the distributions of traffic data for weekdays and Sundays have the same shape and spread

H_0: population median Sunday traffic = population median weekday traffic (traffic equally busy on all days)
H_1: population median Sunday traffic < population median weekday traffic (less traffic on Sundays in summer)

Let α=5%, one-tail test.

Assume H_0 to be true

Data	517	519	521	534	545	546	548	549	553	562
Type	W	W	S	W	W	W	S	S	W	S
Rank	1	2	3	4	5	6	7	8	9	10

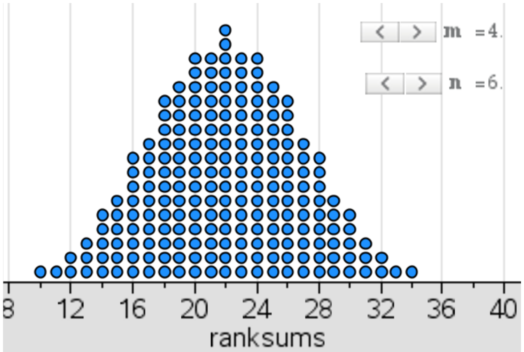
We would reject H_0 if we have a low value of W_s ie lots of Sundays with low traffic count, compared to weekdays. And thus if we have a high value of W_s, then we do not reject H_o.

W_s = sum of ranks of Sunday data m=4
W_w = sum of ranks of Weekday data n=6

So W_s =28 and W_w =27
Minimum rank sum of W_s = $(^1/_2)\times 4\times 5=10$

As m<n, we use W_s

We are interested in $P(W_s \leq 28)$
From tables, for m=4, n=6….
5% critical value is 13



so we are <u>not</u> in the critical region as $W_s = 28 > 13$, which is equivalent to $P(W_s \leq 28) >> 0.05$

So we do no reject H_0 at the 5% level and we conclude that we do not have evidence to suggest that population median Sunday traffic is not equal to the population median weekday traffic, in Summer.

Question 2

From a class of 15 students, 8 are selected at random to receive a daily vitamin pill, the other 7 serve as controls. At the 5% level, can the vitamins be said to have reduced significantly the number of school days missed on account of illness, if the number of days missed are as follows:

Treated	0	2	3	7	8	10	13	18
Control	4	11	12	15	20	21	27	

Solution 2

Justification for using Mann-Whitney: non-paired data
Assumption: distributions of number of days missed have the same shape and spread

H_0: population median treated = population median control (no difference in the number of days missed)
H_1: population median treated < population median control (pills reduce the number of days missed)

Let α=5%, one-tail test.

Assume H_0 to be true

Data	0	2	3	4	7	8	10	11	12	13	15	18	20	21	27
Type	T	T	T	C	T	T	T	C	C	T	C	T	C	C	C
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

We would reject H_0 if we have a low value of W_T i.e. taking the pills means fewer days off
Or we would reject H_0 if we have a high value of W_C

W_C = sum of ranks of control data m=7
W_T = sum of ranks of treated data n=8

So W_C =74 and W_T =46

As m<n, we use W_C

Minimum rank sum of W_C = 1+2+…+6+7=28
Maximum rank sum of W_C = 9+10+…+14+15=84

We are interested in $P(W_C \geq 74)$
Sketch a diagram of the distribution of W_C, showing the minimum and maximum values, and where 74 is on the diagram.
By symmetry of the distribution, $P(W_C \geq 74) = P(W_C \leq 38)$

From tables, for m=7, n=8….
5% critical value is 41

So we conclude that we are in the critical region, as $W_C = 38 \leq 41$ the 5% critical value.
This is equivalent to saying that $P(W_C \leq 38) < 0.05$

Hence, we have evidence to reject H_0 at the 5% level and conclude that we have evidence to suggest that the population median number of days missed for the pupils who took vitamin pills is less than the population median number of days missed for the other pupils. It suggests that the use of vitamin pills does appear to reduce the number of days missed.

Question 3

A car hire firm with 10 identical new cars had 6 fitted with tyres of type A and 4 fitted with tyres of type B. The distances covered, to the nearest 500 miles, before the tread of the front tyres reached the legal minimum were as follows:

Type A	17000	17500	12500	15000	14000	15500
Type B	18000	13500	19500	19000		

Does tyre B last longer than tyre A?

Solution 3

Justification for using Mann-Whitney: non-paired data
Assumption: distributions of number of miles have the same shape and spread

H_0: population median B = population median A (tyres last equally long)
H_1: population median B > population median A (type B tyres last longer)

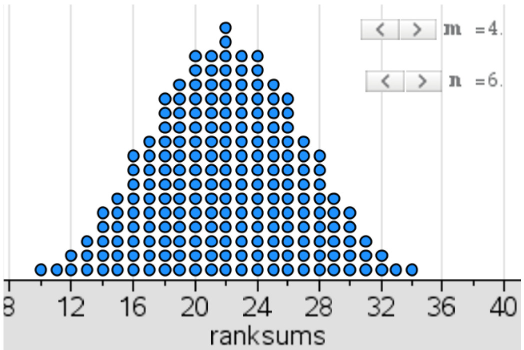
Let α=5%, one-tail test.

Assume H_0 to be true

Data	12500	13500	14000	15000	15500	17000	17500	18000	19000	19500
Type	A	B	A	A	A	A	A	B	B	B
Rank	1	2	3	4	5	6	7	8	9	10

We would reject H_0 if we have low values of W_A i.e. type A tyres don't last as long.
Alternatively, we reject H_o if we have high values of W_B ie type B tyres last longer.

W_A = sum of ranks of tyre A data n=6
W_B = sum of ranks of tyre B data m=4



So W_A =26 and W_B =29
As m<n, we use W_B

Minimum rank sum of W_B = 1+2+3+4 = 10
Maximum rank sum of W_B = 7+8+9+10 = 34

We are interested in $P(W_B \geq 29)$
By symmetry of the distribution, $P(W_B \geq 29) = P(W_B \leq 15)$

From tables, for m=4, n=6….
5% critical value is 13

So $P(W_B \geq 29) = P(W_B \leq 15)$ > 5%, and so we are not in the 5% critical region

Hence, we do not reject H_0 at the 5% level and thus we do not have evidence to suggest that the population medians for the two types of tyres are not equal.
This suggests that the two different types of tyre are equally long lasting.

Question 4

4. A standard memory recall test was administered to a group of S6 pupils and to a group of over forties. The number of items correctly recalled were as follows:

Over 40s	8	11	3	5	9	0	10	
S6 pupils	7	4	10	6	10	2	12	11

Is the memory of the over forty group worse than the S6 pupils?

Solution 4

Justification for using Mann-Whitney: non-paired data
Assumption: distributions of memory recall have the same shape and spread.

H_0: population median over 40's = population median S6 (memory recall equally good)
H_1: population median over 40's < population median S6 (S6 recall is better than over 40's)

Let α=5%, one-tail test.

Assume H_0 to be true

Data	0	2	3	4	5	6	7	8	9	10	10	10	11	11	12
Type	OF	S6	OF	S6	OF	S6	S6	OF	OF	S6	OF	S6	OF	S6	S6
Rank	1	2	3	4	5	6	7	8	9	~~10~~	~~11~~	~~12~~	~~13~~	~~14~~	15
										11	11	11	13.5	13.5	

Note the tied ranks for the numbers 10 and 11. For each set of tied data, the mean rank sum is allocated to each data value. ie. the mean of 10, 11 and 12 is 11. The mean of 13 and 14 is 13.5

High rank sums means they remember more & low rank sums means they remember less.
We would reject H_0 if we have a low value of W_{OF} , or a high value of W_{S6}

W_{OF} = sum of ranks of over 40's data m=7
W_{S6} = sum of ranks of S6 data n=8

So W_{OF} =50.5 and W_{S6} =69.5

As m<n, we use W_{OF}

Minimum rank sum of W_{OF} = 1+2+…+6+7=28
Maximum rank sum of W_{OF} = 9+10+…+14+15=84

We are interested in $P(W_{OF} \leq 50.5)$

From tables, for m=7, n=8….
5% critical value is 41

So we are <u>not</u> in the critical region as W_{OF} = 50.5 > 41.
This is equivalent to saying that $P(W_{OF} \leq 50.5) > 5\%$

Hence, we do not reject H_0 at the 5% level and conclude that we do not have evidence to suggest that the population median numbers of items correctly recalled for both groups are not equal.
This suggests that the memory recall of both groups is equally good.