Wilcoxon Rank-Sum Test

2 independent groups
inference about population medians or distribution

The Wilcoxon Rank-Sum Test (also known as the Mann-Whitney Test in Minitab) is a nonparametric alternative to the 2 independent means t-test. Suppose we have samples from 2 populations, say *A* and *B*. The goal of the Wilcoxon test is to see if *A* and *B* are identically distributed. Because *A* and *B* might be skewed, it is better to use the **median** as the measure of center to determine if their distributions are roughly the same.

Hypotheses for Wilcoxon Rank-Sum Test:

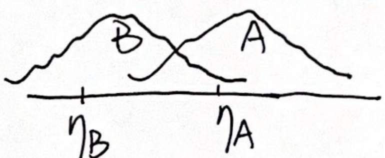
η = Population median (unknown)

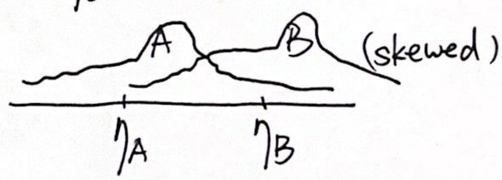
H_0: $A = B$ (distributions are identical)

H_a: $A \neq B$

$A > B$

$A < B$





H_0: $\eta_A = \eta_B$

H_a: $\eta_A \neq \eta_B$, $>$, $<$

Wilcoxon Test Statistic:

- Rank ALL obs from smallest to largest
- Sum ranks for each group: W_A W_B
- TS = W_A = sum of ranks for 1st group.

A	B
• (1)	
	• (2)
	• (3)
• (4)	
• (5)	
• (6)	
	• (7)
⋮	

$$\boxed{U = n_1 n_2 + \frac{n_2(n_2-1)}{2} - W_2}$$
Not required in this class!

p-value Conclusions:

Look at output produced by computer

p-val = prob. of observing data as extreme or more extreme as what we saw if H_0 is true, i.e., $A = B$.

small p-val $\rightarrow$ Rej H_0 $\rightarrow$ Sig Diff in distributions (or population medians)

Example: Suppose we want to compare the grades of two students in each of the following examples. Suppose we want to test if Bob and Jim both have the same grades, and we want to test if Ann's grades are better than Jill's.

Bob	Jim
65 (3)	63 (2)
78 (7)	69 (5)
68 (4)	70 (6)
61 (1)	

Ann	Jill
70 (6.5)	68 (3.5)
68 (3.5)	63 (1)
72 (8.5)	69 (5)
72 (8.5)	65 (2)
70 (6.5)	

$W_B = 3+7+4+1 = 15$ $W_J = 2+5+6 = 13$

- Assumptions:

SRS of grades for each student

→ Are grades representative of Bob's and Jim's grades?

→ Are grades randomly selected?

- Hypotheses:

$H_0: \eta_B = \eta_J$

$H_a: \eta_B \neq \eta_J$

η = median grade for each student in population of grades for each one.

- TS: $W_{Bob} = 15$ (sum of ranks of 1st group)
- p-value: 0.8597 (From output)

→ Not rej H_0 (Fail to rej)

- Conclusions:

94.8% CI for $\eta_1 - \eta_2$:

(−, +) No SIG DIFF

We find No Sig Diff in Bob's & Jim's median scores / distributions of their scores.

assuming the scores in sample are representative of all Bob + Jim scores.

How to deal with the tie? Take the rank mean.

CHECK (made no mistakes when breaking the ties)

$6.5+3.5+8.5+8.5+6.5 = 33.5 \rightarrow W_A$

$3.5+1+5+2 = 11.5 \rightarrow W_J$

$W_A + W_J = 45 = 1+2+\cdots+9 = \frac{(9+1)\times 9}{2} = 45$ ✓

- Assumptions: SAME
- Hypotheses: $H_0: \eta_A = \eta_J$

$H_a: \eta_A > \eta_J$

- TS: $W_A = 33.5$
- p-val: 0.0236 (Adj for ties)

$(1 - 0.0236 = 97.64\%)$

96.3% CI (−, +) → NO SIG DIFF

2-sided CI does NOT have to agree with 1-sided Sig Test.

- Conclusions: Pretty strong evidence to say Ann's grades are better than Jill's assuming those scores are representative of each student.

Mann-Whitney Test and CI: Bob, Jim

```
      N  Median
Bob   4   66.50
Jim   3   69.00

Point estimate for ETA1-ETA2 is -1.50
94.8 Percent CI for ETA1-ETA2 is (-9.00,15.00)
W = 15.0
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.8597
```

Mann-Whitney Test and CI: Ann, Jill

```
       N  Median
Ann    5  70.000
Jill   4  66.500

Point estimate for ETA1-ETA2 is 4.000
96.3 Percent CI for ETA1-ETA2 is (-0.001,8.999)
W = 33.5
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0250
The test is significant at 0.0236 (adjusted for ties)
```

Based on permutations of RANKS

More about p-value: Bob and Jim example

List all possible rankings:

1	4
2	5
3	6
	7

1	3
2	5
4	6
	7

1	3
2	4
5	6
	7

1	3
2	4
6	5
	7

1	3
2	4
7	5
	6

1	2
3	5
4	6
	7

1	2
3	4
5	6
	7

1	2
3	4
6	5
	7

1	2
3	4
7	5
	6

1	2
4	3
5	6
	7

1	
4	
6	

1	
4	
7	

1	
5	
6	

1	
5	
7	

1	
6	
7	

... etc

$$\binom{7}{3} = \frac{7\times 6\times 5}{3\times 2\times 1} = 35$$

$$P(\text{one outcome}) = \frac{1}{35}$$

Only the first column matters

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2
	2	2	2	2	2	3	3	3	3	4	4	4	5	5	6	3	3
	3	4	5	6	7	4	5	6	7	5	6	7	6	7	7	4	5
W	6	7	8	9	10	8	9	10	11	10	11	12	12	13	14	9	10
W ordered	6	7	8	8	9	9	9	10	10	10	10	11	11	11	11	12	12

	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	5
	3	3	4	4	4	5	5	6	4	4	4	5	5	6	5	5	6	6
	6	7	5	6	7	6	7	7	5	6	7	6	7	7	6	7	7	7
W	11	12	11	12	13	13	14	15	12	13	14	14	15	16	15	16	17	18
W ordered	12	12	12	12	13	13	13	13	14	14	14	14	15	15	15	16	16	17	18

35 possibilities

15 of which have sum of ranks for smallest group of 13 or more

p-value = Prob of getting a test statistic as extreme as what we got if null hypothesis is true (2 sided test)

= Prob (W >= 13) x2

= 15/35 x 2

= 0.4286 x 2

= 0.8571

Extra Credit 3

(0.5 pts)

* Optional * Challenging

Data:

	Bob	Jim
	65	63
	78	69
	68	70
	61	71

1. Mimicing the way we did on the Page 56 of finding the p-value, find the p-value of Wilcoxon Rank-Sum Test with the data given above.

2. Find the p-value for Ann/Jill example on Page 54. (You shall be able to find a p-value that is very close to 0.0236, but NOT exactly equal. Why is that?)

Deadline: 3/18/2023 11:59pm

Upload a pdf file on Canvas "Extra Credit 3"

[No email ; No paper]