



The Mann-Whitney U Test

DESMOND AYIM-ABOAGYE, PHD

Ordinal Data

- ▶ Usually interested in how participants rank order some set of stimuli within the context of an experiment.
- ▶ How can ordinal or “ranked” data be analyzed?
- ▶ Ordinal data cannot be analyzed by the chi-square or any of the other inferential (parametric) tests we examined.
- ▶ Three tests to test hypotheses with ordinal data:
 - ▶ 1. The Man-Whitney U test,
 - ▶ 2. The Wilcoxon signed-ranks test
 - ▶ 3. The Spearman Correlation Coefficient, or Spearman r .

The Mann-Whitney U Test

- ▶ **The Mann-Whitney U Test** is a nonparametric statistic used to identify a difference between two independent samples of rank-ordered (ordinal) data.
- ▶ Statistical Assumptions underlying this test:
 - ▶ * The data are based on an ordinal scale of measurement
 - ▶ * The observations were drawn or selected independently of one another.
 - ▶ * There are no "ties" (i.e., same values with different ranks) between rankings (ties do occur, however, and a quick procedure for dealing with them is presented in Data Box 14.D When the majority of ranks in a data set are tied, however, consult statistical works like Hays (1988) or Kirk (1990) for guidance.

EXAMPLE:

- ▶ Perhaps a linguist is interested in comparing the effectiveness of traditional, classroom-based language learning versus total immersion learning where elementary students are concerned. The linguist randomly assigns a group of 18 fourth-graders to either a traditional Spanish language class (i.e., **the teacher gives directions in English, though the emphasis is on learning to speak Spanish**) or a total immersion class (i.e., **the teacher speaks exclusively in Spanish**). At the end of the school year, a panel of judges gives an age-appropriate Spanish-language test to the students, subsequently using the scores to rank the children's linguistic skills from 0 to 100 (the judges remain unaware of which learning technique each child was exposed to). The rankings were then categorized by the respective teaching techniques the students were exposed to (see Table 14.4).

Hypotheses: Null versus Alternative

- ▶ H_0 : There will be no systematic difference between the Spanish-speaking skills of the traditional-learning group and the total immersion group.
- ▶ H_1 : There will be systematic difference between the Spanish-speaking skills of the traditional-learning group and the total immersion group.

Tables 14.4 Spanish-Speaking Skills Resulting from Linguistic Pedagogy

Traditional classroom (English & Spanish spoken)	Total Immersion (Spanish only spoken)
35	75
56	83
42	77
78	92
82	85
72	95
62	73
42	83
51	
38	

Note: Each number represents the relative ranking of a student's ability to speak Spanish after 1 year of receiving one mode of instruction.

Table 14.5. Combined Ranks for Spanish-Speaking Skills Resulting from Linguistic Pedagogy

1 Ordered Raw Scores of Two Groups	2 Ranks of Scores of Two Groups	3 Group Identification	4 Ranks for Group A	5 Ranks for Group B
35	1	A	1	
38	2	A	2	
42	3.5	A	3.5	
42	3.5	A	3.5	
51	5	A	5	
56	6	A	6	
62	7	A	7	
72	8	A	8	
73	9	B		9
75	10	B		10
77	11	B		11
78	12	A	12	
82	13	A	13	
83	14	B		14
85	15	B		15

Handling Tied Ranks in Ordinal Data

- ▶ Rank of tied scores = $\frac{\text{sum of rank positions by tied scores}}{\text{number of tied scores present}}$

- ▶ Rank of tied Scores $\frac{3 + 4}{2}$

- ▶ Rank of tied Scores = $7/2$

- ▶ Rank of tied Scores = 3.5

- ▶ Rank of tied Scores = $5 + 6 + 7 / 3$

- ▶ Rank of tied Scores = $18/3$

- ▶ Rank of tied Scores = 6

Formula U statistic:

▶ $U_A = N_A N_B + \frac{N_A (N_A + 1)}{2} - \sum R_A$.

▶ $\frac{2}{2}$

▶ $U_A = (10)(8) + \frac{10(10+1)}{2} - 61$

▶ $\frac{2}{2}$

▶ $U_A = 135 - 61$

▶ $U_A = 74$.

Formula for UB is:

▶ $UB = N_A N_B + \frac{N_B (N_B + 1)}{2} - \sum RB$

▶ $\frac{2}{2}$

▶ $UB = (10)(8) + \frac{8(8+1)}{2} - 110$

▶ $\frac{2}{2}$

▶ $UB = 80 + 36 - 110$

▶ $UB = 116 - 110$

▶ $UB = 6.$

Rejection or Acceptance of Hypothesis (see **Table B.8** in **Appendix B**)

- ▶ To select the $U_{critical}$ value for Mann-Whitney U_{test}
- ▶ Sample sizes of the groups must be known A ($N_A = 10$) & B ($N_B = 8$)
- ▶ Significance level of step 2 (i.e., .05)
- ▶ **To be significant, the smaller computed U must be equal to or less than critical U.**
- ▶ To determine whether we can reject H_0 , we compare the two values U_A (74) and U_B (6) against the $U_{critical}$ value 17.
- ▶ Because the U_B of 6 is *less than* the $U_{critical}$ value of 17, we reject H_0 .

Reject H_0

- ▶ The two groups of ranks represent different populations, such that students in language immersion group had higher language proficiency rankings than those students who learned in the traditional manner A and B in columns 4 and 5, respectively. Table 14.5.
- ▶ Which means the language immersion group demonstrated relatively greater proficiency speaking Spanish *than* the traditional learning group.