

Reflection #9. This reflection focuses on nonparametric tests. Some of the questions require analysis of the data set using both nonparametric and its parametric counterpart. In such cases, pay particular attention to assumption violation to see how much influence if any it has on test results.

1. Comparison of parametric and nonparametric tests for paired samples

The Allan Course Preparation is designed to help students achieve better scores on SAT exams. The table below lists results for randomly selected students.

Students	A	B	C	D	E	F	G	H	I	J
SAT score before course	700	840	830	860	840	690	830	1180	930	1070
SAT score after course	800	840	820	980	980	800	800	1270	1080	1220

Use the following tests to test the hypothesis that the course has no effect on SAT scores at $\alpha = 0.05$:

a. Sign test. **Sign Test on Paired Data**

Null Hypothesis: The SAT prep course does not have any effect on the students' SAT scores.

Alternative: The SAT prep course does have an effect on the students' SAT scores.

We are conducting a two-tailed test.

Students	A	B	C	D	E	F	G	H	I	J
SAT score before course	700	840	830	860	840	690	830	1180	930	1070
SAT score after course	800	840	820	980	980	800	800	1270	1080	1220
Difference	-100	0	10	-120	-140	-110	30	-90	-150	-150
Sign	-	Ignore	+	-	-	-	+	-	-	-

Because $n \leq 25$, our test statistic, $X = 2$, since we only have 2 positive signs appear.

There are $n = 9$ nonzero differences, and $\alpha = 0.05$. From Table I, the critical value is 1.

Because $X > critical\ value$, we fail to reject the null hypothesis. Thus, there is not enough evidence to support the claim that the SAT prep course does have an effect on student SAT scores.

b. Wilcoxon signed-ranks test.

Null Hypothesis: The SAT prep course does not have any effect on the students' SAT scores.

Alternative: The SAT prep course does have an effect on the students' SAT scores.

Students	A	B	C	D	E	F	G	H	I	J
SAT score before course	700	840	830	860	840	690	830	1180	930	1070
SAT score after course	800	840	820	980	980	800	800	1270	1080	1220
Before-After	-100	0	10	-120	-140	-110	30	-90	-150	-150
 Before-After 	100	0	10	120	140	110	30	90	150	150
Rank	4		1	6	7	5	2	3	8.5	8.5

Signed Rank	-4		1	-6	-7	-5	2	-3	-8.5	-8.5
--------------------	----	--	---	----	----	----	---	----	------	------

Negative Sum Rank: $|(-4) + (-6) + (-7) + (-5) + (-3) + (-8.5) + (-8.5)| = |-42|$

Positive Sum Rank: $|1 + 2| = |3|$

Test statistic: 3

There are $n = 9$ nonzero differences, and $\alpha = 0.05$. From Table J, the critical value is 6.

Because $3 \leq 6$, we reject the null hypothesis. Thus, there is enough evidence to support the claim that the SAT prep course does have an effect on student SAT scores.

c. t test for a claim about two dependent samples.

Null Hypothesis: The SAT prep course does not have any effect on the students' SAT scores.

Alternative: The SAT prep course does have an effect on the students' SAT scores.

For this test, we must assume that the differences are approximately normally distributed. However, when we plot a histogram of the differences, the data does not take on a normal shape.

The test statistic is calculated by the formula, $t_0 = \frac{\bar{x}_d}{s_d/\sqrt{n}}$, where \bar{x}_d = The mean value of the sample of differences, s_d = the standard deviation of the sample of differences, and n = number of pairs.

Students	A	B	C	D	E	F	G	H	I	J
SAT score before course	700	840	830	860	840	690	830	1180	930	1070
SAT score after course	800	840	820	980	980	800	800	1270	1080	1220
Before-After	-100	0	10	-120	-140	-110	30	-90	-150	-150

$$n = 10$$

$$\bar{x}_d = -82$$

$$s_d = \sqrt{\frac{\sum(d_i - \bar{x}_d)^2}{n - 1}} = \sqrt{\frac{42960}{9}} = 69.09$$

Thus, our test statistic is:

$$t_0 = \frac{-82}{69.09/\sqrt{10}} = \frac{-82}{21.8482} \approx -3.753$$

Because we are working with a two-sided hypothesis, we will reject H_0 if $t_0 < -2.262$ or $t_0 > 2.262$.

Because $-3.753 < -2.262$, we reject the null hypothesis. Thus, there is enough evidence to support the claim that the SAT prep course does have an effect on student SAT scores.

d. How do the results of these tests support the statement that nonparametric tests lack the sensitivity of parametric test?

The Sign test failed to reject the null hypothesis where the other two tests did. This does demonstrate that non-parametric tests can be less sensitive to changes in data. We also did violate an assumption in the parametric test, which didn't seem to matter as we got the same value as our Wilcoxon Signed-Ranks Test.

2. The Mungo Opinion Research Organization assigned a pollster to collect data from 30 randomly selected adults. As the data were submitted to the company, the gender of each interviewed participant was noted. The sequence obtained is shown in the accompanying list.

M M F M M F M M F M F M M F M
 F M F M M F M M F M M F M M M

- a. At the 0.05 significance level, test the claim that the sequence is random.

Null Hypothesis: Pollsters were selected at random, according to gender

Alternative Hypothesis: The null hypothesis is not true.

Test Statistic: First, we must find the number of runs.

<u>RUN</u>	1	2	3	4	5	6	7	8	9	10	11
<u>GENDER</u>	MM	F	MM	F	MM	F	M	F	MM	F	M
<u>RUN</u>	12	13	14	15	16	17	18	19	20	21	
<u>GENDER</u>	F	M	F	MM	F	MM	F	MM	F	MMM	

There are 21 runs, with 20 males (m) and 10 females (n).

Because $n \leq 20$, $m \leq 20$, and $\alpha = 0.05$, we will use Table M to find the critical values. So, our critical values are 9 and 20. Because $21 > 20$, we will reject the null hypothesis. Thus, there is significant evidence that pollsters were not selected at random according to gender.

- b. At the 0.05 significance level, test the claim that the proportion of women is different from 0.5

Null Hypothesis: The proportion of women is equal to 0.5; $p = 0.5$

Alternative Hypothesis: The proportion of women is not equal to 0.5; $p \neq 0.5$

Our test statistic is given by $z_0 = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}}$, where $\sigma_{\hat{p}} = \sqrt{\frac{p_0(1-p_0)}{n}}$, $p_0 = 0.5$, and $\hat{p} = \frac{10}{30} = 0.333$.

$$\sigma_{\hat{p}} = \sqrt{\frac{0.5(1-0.5)}{30}} = \sqrt{\frac{.25}{30}} \approx .0913$$

$$z_0 = \frac{0.333 - 0.5}{.0913} = \frac{-0.1667}{.0913} \approx -1.826$$

Because this is a two-tailed test, our P-value is calculated as

P-value = $2 * P(z \geq |z_0|) = 2 * P(z \geq 1.826) = 2 * 0.0339 = 0.0678$. Because our P-value is greater than or equal to α , we fail to reject the null hypothesis. Thus, there is not sufficient evidence to conclude that the proportion of women deviates from 0.5.

- c. What do the preceding results suggest?

The run test suggests that the sequence in which genders were polled was not random. The proportion test tells us that there wasn't enough evidence to conclude the proportion of women deviates from 50%. However, because the run test suggests that the data was not obtained via a random sample, we may have violated an assumption for the population proportion test.

Is the sample biased against women?

The bias does not appear to be against women specifically but could be in the order of selection.

Was the sample obtained in a random sequence?

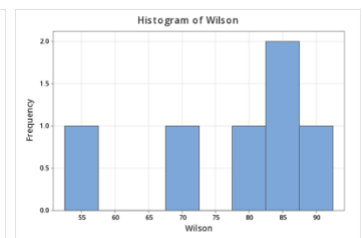
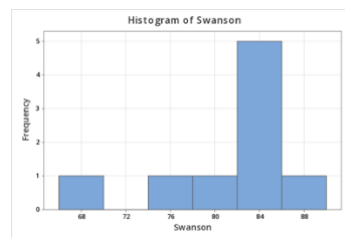
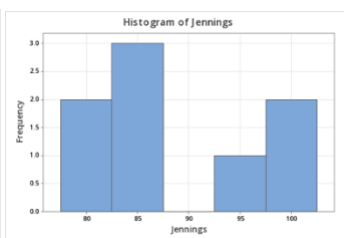
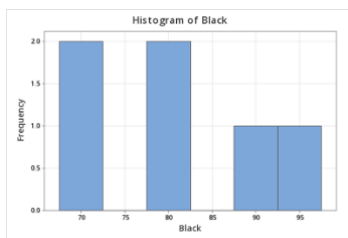
No the run test tells us there was a strong possibility that the sample was not obtained randomly.

If you are the manager, will you have any problems with these results?

Yes, sampling should occur randomly, so as a manager I would revisit how we are selecting people we poll.

3. Course evaluation ratings for four college instructors are shown in the following table. Use $\alpha = .05$ and test for a significant difference among the rating for these instructors. What is your conclusion?

Instructor			
Black	Jennings	Swanson	Wilson
88	87	88	80
80	78	76	85
79	82	68	56
68	85	82	71
96	99	85	89
69	99	82	87
	85	84	
	94	83	
		81	



Based on the graphs of the data, it is clear that this is not normally distributed, so we must use the Kruskal-Wallis Test.

Null Hypothesis: There is no difference in the ratings for these instructors

Alternative Hypothesis: There is a difference in the ratings for these instructors.

The Critical value is found by $\chi^2_{3,0.05} = 7.815$. We know that the degrees of freedom is 3, because there are 4 groups.

We know that $N = 29$, because there are 29 ratings.

Instructor							
Black		Jennings		Swanson		Wilson	
Course Eval Rating	Rank	Course Eval Rating	Rank	Course Eval Rating	Rank	Course Eval Rating	Rank
88	23.5	87	21.5	88	23.5	80	9.5
80	9.5	78	7	76	6	85	18.5
79	8	82	13	68	2.5	56	1
68	2.5	85	18.5	82	13	71	5
96	27	99	28.5	85	18.5	89	25
69	4	99	28.5	82	13	87	21.5
		85	18.5	84	16		
		94	26	83	15		
				81	11		
Rank Sum	74.5		161.5		118.5		80.5

To find the test statistic, we will use the formula:

$$\begin{aligned}
 H &= \frac{12}{N(N+1)} \left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} + \frac{R_4^2}{n_4} \right) - 3(N+1) \\
 &= \frac{12}{29(30)} \left(\frac{74.5^2}{6} + \frac{161.5^2}{8} + \frac{118.5^2}{9} + \frac{80.5^2}{6} \right) - 3(30) \\
 &\approx 4.146
 \end{aligned}$$

Because the test statistic is less than the critical value ($4.146 < 7.815$), we fail to reject the null hypothesis. Thus, there is not sufficient evidence to conclude that there is a difference in the ratings for the instructors.

4. The following data show the rankings of 11 states based on expenditure per student (ranked 1 highest to 11 lowest) and student-teacher ratio (ranked 1 lowest to 11 highest).

State	Expenditure per Student	Student-Teacher Ratio
Arizona	9	10
Colorado	5	8
Florida	4	6
Idaho	2	11
Iowa	6	4
Louisiana	11	3
Massachusetts	1	1
Nebraska	7	2
North Dakota	8	7
South Dakota	10	5
Washington	3	9

- a) Compute the rank correlation between expenditure per student and student-teacher ratio?

Null hypothesis: $\rho = 0$ (There is no correlation between the two variables)

Alternative hypothesis: $\rho \neq 0$ (There is a correlation between the two variables)

State	Expenditure per Student	Student-Teacher Ratio	d_i^2
Arizona	9	10	1
Colorado	5	8	9
Florida	4	6	4
Idaho	2	11	81
Iowa	6	4	4
Louisiana	11	3	64
Massachusetts	1	1	0
Nebraska	7	2	25
North Dakota	8	7	1
South Dakota	10	5	25
Washington	3	9	36
		SUM	250

Now, we find the r_s statistic by using the formula:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} = 1 - \frac{6(250)}{11(120)} = 1 - \frac{1500}{1320} = 1 - 1.13636 \approx -0.1364$$

- b) At $\alpha = .05$, determine whether there is a relationship between expenditure per student and student-teacher ratio?

Using table L for $\alpha = 0.5$, where $n = 11$ pairs of values, the critical value is 0.623. Because this is for a two-tailed test, we reject if $r_s < -0.623$ or $r_s > 0.623$. Because $-0.1364 > -0.623$, we fail to reject the null hypothesis. Thus, there is no significant relationship between expenditure per student and student-teacher ratio.

5. The gap between the earnings of men and women with equal education is narrowing but not closed. Sample data for seven men and seven women with bachelor's degrees are as follows. Data are shown in thousands of dollars.

Men	Women
35.6	49.5
80.5	40.4
50.2	32.9
67.2	45.5
43.2	30.8
54.9	52.5
60.3	29.8

- a) What is the median salary for men? For women?

The median salary for men is \$54.9 thousand, while the median salary for women is \$40.4 thousand

- b) Use $\alpha = .05$ and conduct the hypothesis test of identical population distribution. What is your conclusion?

For this problem, we are using the Wilcoxon Rank-Sum Test.

Null Hypothesis: There is no difference in the median salary for men vs. women (median salary for men = median salary for women)

Alternative hypothesis: This is a difference in the median salary for men vs. women (median salary for men \neq median salary for women)

Because $n_1 = 7 = n_2 \leq 10$, we will find the test statistic $T_x =$ rank sum of the sample with the fewest members, but because they are the same, we can use either one.

Men	RANK	Women	RANK
35.6	4	49.5	8
80.5	14	40.4	5
50.2	9	32.9	3
67.2	13	45.5	7
43.2	6	30.8	2
54.9	11	52.5	10
60.3	12	29.8	1
SUM	69	SUM	36

So, $T = 69$ or $T = 36$.

Because $n \leq 10$, we find the critical value by using table K. Since the hypothesis is two-sides, we will either reject the hypothesis if $T \leq T_L$ or $T \geq T_U$. Since $n_1 = n_2 = 7$, $\alpha = 0.5$, using table K, we know that $T_L = 37$ and $T_U = 68$.

Using $T = 36 \leq 37$, or $T = 69 \geq 68$, we will reject the null hypothesis. Thus, there is enough evidence to conclude that the median salary for men is different than the median salary for women.

c) What assumptions did you make in performing the test in (b)?

We are assuming that the data can be ranked (salary), and also, that the two samples were selected in both an independent and random fashion.