Extra Topic! (Chapter 4): Alternatives to the t-Tools (Distribution Free Methods)

Motivation: When situations occur where the assumptions for the conventional t-tools (parametric methods) are grossly violated, and no useful transformation can be found, we must consider alternative methods to analyze the data. The methods presented in this chapter are called distribution-free methods (or non-parametric methods) because there is no assumption needed for the distribution of the populations involved.

We will look at two alternative methods to the t-tools:

1) one for two independent samples (Rank-Sum Test)
2) one for paired data (Wilcoxon Signed-Rank Test)

These methods are particularly useful when:
- **outliers** are present,
- when **sample sizes are small**, or
- when ***censoring*** occurs.

When ***censored*** observations or outliers exist it is often better to consider a non-parametric approach because non-parametric tests are resistant to outliers. Non-parametric approaches incorporate rank transformations. Depending on the method, the observations are only considered in their ranks. Thus, the largest value in the data set will have rank *n* regardless of how large it is.

Consider the following sample of times (in minutes) to complete a task:

10, 20, 30, 40, 60+

The plus sign indicates that this observation was censored. That is, the study ended after 1 hour. For this individual all we know is that it took at least 60 minutes. Thus, we have no known values to consider if we were to calculate the average. However, applying a rank transformation, this value would take on rank 5 regardless.

Example: Who's faster?

By analyzing the results of, say, the 800m final results of the 2008 Beijing Olympics, we might be able to determine, once and for all, which gender is faster.

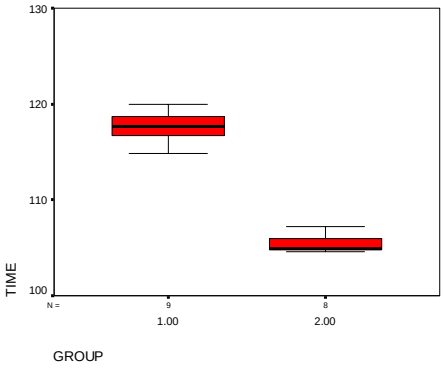
Rank Transformation: This is for illustration only; you will never have to do this by hand.

Time	Nation	Group	Order	Rank
104.65	Kenya	Male	1	1
104.70	Sudan	Male	2	2
104.82	Kenya	Male	3	3
104.94	Canada	Male	4	4.5
104.94	Bahrain	Male	5	4.5
105.88	Cuba	Male	6	6
105.96	Algeria	Male	7	7
107.19	Algeria	Male	8	8
114.87	Kenya	Female	9	**9**
116.07	Kenya	Female	10	**10**
116.73	Morocco	Female	11	**11**
116.94	Russian Fed.	Female	12	**12**
117.68	Mozambique	Female	13	**13**
118.24	Jamaica	Female	14	**14**
118.73	Ukraine	Female	15	**15**
120	Russian Fed.	Female	16	**16**
120	Italy	Female	17	**16**

Report

TIME

GROUP	Mean	N	Std. Deviation
1.00	117.6956	9	1.7363
2.00	105.3850	8	.8911
Total	111.9024	17	6.4785



Preliminary Analysis

It appears from the averages that the male athletes are faster, but is there "significant" evidence to draw the conclusion?

Comments on Distribution Assumptions

Although the distribution of females seems normal, the distribution of males seems a bit right-skewed. The variability appears to be roughly similar in terms of the interquartile ranges and the ratio of the standard deviations is about 1.95.

Nevertheless, the last two observations appear to be ***censored***, since all we know is that it took at least 120 seconds (or two minutes) to cross the finish line. A non-parametric approach may be better.

Procedure, Statistical Results, and Conclusions

Because we have two independent samples, the rank-sum procedure (an alternative to the two-independent sample t-test) is necessary.

Descriptive Statistics

	N	Sum	Mean	Std. Deviation
RANK	17	152.00	8.9412	4.9525
Valid N (listwise)	17			

Test Statistics[b]

	TIME
Mann-Whitney U	.000
Wilcoxon W	36.000
Z	-3.468
Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.000[a]

a. Not corrected for ties.

b. Grouping Variable: GROUP

Ranks

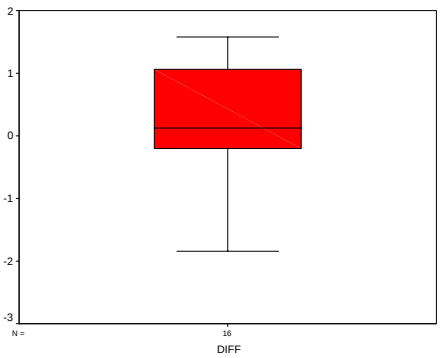
	GROUP	N	Mean Rank	Sum of Ranks
TIME	1.00	9	13.00	117.00
	2.00	8	4.50	36.00
	Total	17		

For the final 800m races, it appears that the male athletes are significantly faster due to strong to convincing evidence against the null hypothesis (p-value < 0.0005).

Example: Maybe instead we should compare two times from each racer.

Descriptive Statistics

	N	Sum	Mean	Std. Deviation
DIFF	16	2.76	.1725	1.0731
Valid N (listwise)	16			



Comments on Distribution Assumption

Although the assumptions for the use of t-tools are not *seriously* violated (though noticeable right-skewness exists), the sample size is somewhat small. As a result, checking the distribution can be difficult.

Procedure, Statistical Results, and Conclusions

Because this is paired data, apply the Wilcoxon – Signed Rank test.

SPSS Output:

Ranks

		N	Mean Rank	Sum of Ranks
TIME2 - TIME1	Negative Ranks	5[a]	9.80	49.00
	Positive Ranks	11[b]	7.91	87.00
	Ties	1[c]		
	Total	17		

a. TIME2 < TIME1

b. TIME2 > TIME1

c. TIME1 = TIME2

Test Statistics[b]

	TIME2 - TIME1
Z	-.982[a]
Asymp. Sig. (2-tailed)	.326

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

Signed-Rank Transformation: This is for illustration only, you will never have to do this by hand.

Athlete	Time 1	Time 2	Diff	Abs Diff	Rank
9	120	120	**0**	0	1
14	104.94	104.95	0.01	0.01	2
6	118.24	118.28	0.04	0.04	3
12	104.82	104.73	**-0.09**	0.09	4
11	104.7	104.91	0.21	0.21	5
16	105.96	105.63	**-0.33**	0.33	6
15	105.88	106.4	0.52	0.52	7
13	104.94	105.85	0.91	0.91	8
5	117.68	118.61	0.93	0.93	9
2	116.07	117.28	1.21	1.21	10
3	116.73	118.03	1.3	1.3	11
4	116.94	118.31	1.37	1.37	12
7	118.73	117.32	**-1.41**	1.41	13
10	104.65	106.23	1.58	1.58	14
17	107.19	105.54	**-1.65**	1.65	15
8	120	118.16	**-1.84**	1.84	16
1	114.87	117.31	2.44	2.44	17