

Do the Right Test! A Nonparametric Approach

Brian E. Smith
McGill University

Presented at the USCOTS conference
Columbus, Ohio
May 19 – 21, 2005

Abstract:

For my second course in Statistics I emphasize the importance of non-parametric tests. I try to impress on my students the dangers of "doing the wrong" test and tie this in with the importance of goodness of fit tests for a normal distribution. Examples of non-parametric tests will be provided.

Suppose that in a test for the equality of two populations (means or medians) we have obtained the following differences in a sample of 14 paired differences:

-7, 34, -85, 140, 175, 180, 150, 99, 65, -36, 18, -8, 3, 1

When we run a sign test on these data we obtain the following results (Minitab output):

Sign Test for Median: Difference

Sign test of median = 0.00000 versus not = 0.00000

	N	Below	Equal	Above	P	Median
Difference	14	4	0	10	0.1796	26.0

We see that, at the 0.05 level of significance, we cannot reject the hypothesis $H_0: \eta_1 = \eta_2$ (p-value = 0.1796).

The Wilcoxon test yields the following results:

Wilcoxon Signed Rank Test: Difference

Test of median = 0.000000 versus median not = 0.000000

	N	for	Wilcoxon	Estimated	
	N	Test	Statistic	P	Median
Difference	14	14	82.0	0.069	50.00

Once again we cannot reject $H_0: \eta_1 = \eta_2$ at the 0.05 level of significance, but we see that the Wilcoxon test statistics of 82 is approaching significance with a p-value = 0.069.

Finally, when we apply the paired sample t-test we get:

One-Sample T: Difference

Test of $\mu = 0$ vs not = 0

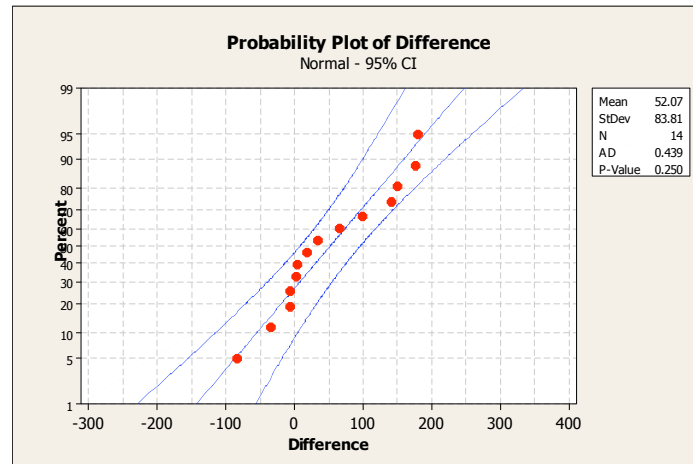
Variable	N	Mean	StDev	SE Mean	95% CI	T	P
Difference	14	52.0714	83.8107	22.3993	(3.6806, 100.4623)	2.32	0.037

In this case we see that the p-value is 0.037 so that we can conclude that there is a significant difference between the population means. We also note the 95% confidence interval is between 3.68 and 100.46, so that the interval does not contain 0.

In this problem it is critical to know which test is the right one, since they yield different conclusions!

To perform the t-test we need to assume a normal distribution. Do the data support this assumption?

The normal probability plot below indicates that the population of differences may be assumed to be normal.



Therefore we are justified in using the paired samples t-test and we can conclude that there is a significant difference between the population means. Note that if we had performed the wrong test (either the sign test or the Wilcoxon signed rank test) we would not have rejected the null hypothesis of no difference, because of the lack of power associated with these tests.

We conclude that in general it is preferable to perform parametric tests, and we should only resort to nonparametric tests in those cases where the underlying distributions are not normal, as demonstrated by performing tests for normality (goodness of fit, normal probability plots, etc.)