

Determining if rain lowers the pH of local lake water: An example of paired samples

This worksheet accompanies the article:

Martinez-Dawson, R. (2003) "Incorporating laboratory experiments in an introductory statistics course," *Journal of Statistics Education* [Online], 11(1).
<http://www.amstat.org/publications/jse/v11n1/martinez-dawson.html>

Purpose:

The purpose of this experiment is to illustrate the concepts of sampling involving paired samples, statistical analysis of the samples, and drawing a conclusion based on the analysis. You will also produce a 95% confidence interval for μ_d .

Ecologists have been concerned about the effect of pollution on the environment, specifically acid rain. The lower the pH of a solution, the more acidic it is. Water samples were taken from various sites on local lakes before and just after it rained.

Materials:

You will need the following materials: lake water samples labeled by site and if it was taken before or after it rains, pH meter, stirrer, and kimwipes.

Procedures:

Stir each water sample in the jar. Measure pH using the pH meter and record the pH in the table below. Rinse the electrode in distilled water after each use and dry with a kimwipe. Dry the electrode before placing it in the electrode storage solution.

Site	pH-Before Rain	pH-After Rain

Statistical analysis:

Conduct a hypothesis test to determine for the Clemson area if the mean pH of lake water after it rains is lower than the mean pH of lake water before it rains. ($\alpha = .05$)

1. State the null and alternative hypotheses. Define the parameter(s). (4 pts)
2. Calculate the test statistic. Show all your work. (12 pts)
3. Determine the approximate p-value. Be as precise as possible. (8 pts)
4. What is the decision? Explain. (2 pts)
5. What is the conclusion? (2 pts)

Conduct the hypothesis test using SAS. Type the following code (in bold) to perform the hypothesis test using the p-value method. A description of any new code will appear to the right of the SAS code.

```
data rain;  
input before after;  
cards;
```

Type the pH of the lake water before raining, skip a space, and then type the pH of the lake water after raining.

```
proc ttest alpha=.05;  
paired before*after;  
title Type your name here;  
run;  
quit;
```

Questions based on the SAS output

6. On the printout, what does the value under the "Mean" column represent? Give the symbol. (2 pts)

7. What does the value under the "Std dev" column represent? Give the symbol. (2 pts)
The formula for computing the test statistic is:

$$t_{obs} = \frac{\bar{d} - \delta_0}{s_d / \sqrt{n}}$$

8. Put the values from the printout correctly into the test statistic formula. (4 pts)
 $t_{obs} =$

9. Identify the test statistic **on the printout** by labeling it t_{obs} , and write the degrees of freedom in the space below. (2 pts)

10. Based on the printout, determine the p-value for this hypothesis test. (2 pts)

11. Draw the t-distribution and indicate the area that corresponds to the p-value. (2 pts)

12. What is the decision? Explain. (2 pts)

13. Based on the hypothesis test, what is the conclusion? (2 pts)

14. Give the LCL and UCL for the 95% confidence interval for μ_d . (2 pts)

LCL=

UCL=

15. How would you change this experiment to make it an example of independent samples? (2 pts)