

Math 263: Excel Assignment 6 on Hypothesis Tests: Impact of Carbon Dioxide on Pine Trees

Due Thursday April 3, at the start of class. Hand in a paper printout of the answers to the questions.

Answers that appear to be copied will be treated as an integrity violation.

Biological Background

The quantity of carbon dioxide (CO_2) in the atmosphere has been increasing for the past half century. This assignment looks at data from an experiment investigating the effect of increases in CO_2 on plant growth. Since carbon is a nutrient needed by plants, one possibility is that additional CO_2 increases plant growth. Since plants also absorb CO_2 , it is possible that increased growth will lead to increased carbon absorption.

The data we will use comes from an experiment in which some pine trees in a forest were grown in the circular regions shown with additional CO_2 and some were grown in other circular regions with the usual atmosphere.¹ One measure of their growth is “dbh”, or *diameter at breast height*, in centimeters. The spreadsheet *PineTrees.xlsx* contains the dbh of 644 trees.



¹ <http://www.sciencemag.org/content/292/5514/95.full>, reported by Michael Lavine <https://www.math.umass.edu/~lavine/>

Purpose of Assignment: Decide if CO₂ Increases Plant Growth

You will test the data on dbh and see whether trees grown with additional CO₂ have significantly different values of dbh.

- Treatment groups are Groups 2, 3, 4
- Control groups are Groups 1, 5, 6

1. Before you start, you will need to sort the dbh data into treatment and control groups. One way to do this is to sort by Group # and then copy and paste the Treatment groups into one column and the Control groups into another. Best is to use a new page of the spreadsheet.

For each sample, control and treatment, fill out a table like the following:

	Treatment	Control
Mean		
Standard deviation		
Number in sample		

The commands you will need are:

= AVERAGE(array)
=STDEV.S(array)

2. What is the standard error of the *difference in means* of the two samples?
3. Do a hypothesis test to decide if the Treatment regions have significantly different dbh values than the Control regions. *Lay the reasoning of the test out in detail, showing each of the four steps clearly.*

To find the p-value you can use this or your calculator

=T.DIST(*t*-value, degrees of freedom, true)

4. Find the 95% confidence interval for the difference in means.
5. What is the relation between the confidence interval you just found and the hypothesis test you did earlier? *Explain the reasoning precisely.*