

R Class Homework 9

Assignment

- 1) Complete the 'sig_t' function in R Class Exercise 3. **The function is started and outlined for you in CE3.**
- 2) Get the p-values from sig_t for data columns 'l', 'v', and 'whole'.
 - **Check your answer:** The p-value for 'l' should be 0.4334 (unless I did it wrong).

sig_t function

This function should calculate a t statistic from two groups, and return the two-tailed p-value for that t-statistic given the degrees of freedom. We will use this function to get the p-value for a t-test of significant differences between clay (soil=="C") and sand (soil=="S") groups of specific-leaf-area values. We will do this for observed SLA values for 'l', 'lv', and 'whole' leaf structural groupings.

Later, this function will be used in a bootstrap iteration of t-tests on artificially generated samples with inflated sample sizes. These will be the basis for a 'power analysis' to see how many more SLA samples we're likely to need to distinguish soil types.

The two main equations for the t-test function copied below. Remember that for the sum of squares for each group, you are taking the sum of a vector minus a single mean. The subtraction of one mean from an entire vector of values is a *vectorized* operation. E.g., $c(1,2,3) + 10 = c(11, 12, 13)$.

Handwritten equations for the t-test function:

Pooled variance

$$\frac{\sum (grp1.x_i - \overline{grp1.x})^2 + \sum (grp2.x_i - \overline{grp2.x})^2}{d.f.}$$

SS2

t-statistic

$$t = \frac{\overline{grp1.x} - \overline{grp2.x}}{\sqrt{\text{pooled.var} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Annotations:

- SS2: Sum of Squares
- d.f.: Degrees of freedom
- Group means: $\overline{grp1.x}$ and $\overline{grp2.x}$