**Lab 6** Self Assessment.

The data in arabidop.csv are measurements of the alanine concentration in samples from three genotypes of Arabidopsis. This is part of a gene knockout study, but details of that are suppressed. There are four plants from each of three genotypes. It is appropriate to consider plant the experimental unit even though genotype is not randomly assigned to a plant. Each plant was measured twice (technical rep). Before analysis, you need to calculate the average concentration for each plant. We saw the averages last week. Now you see the original data from all three genotypes.

The variables are the Genotype, the BIOLogical replicate number (i.e. the plant), and the TECHnical replicate number, and the ALAnine concentration measured in that replicate.

Note: In lecture, we’ve only talked (so far) about analyzing two groups. Almost everything we have done works equally well with 3 groups. The exceptions are obvious, e.g., t-tests are only for 2 groups.

1) Calculate the average alanine concentration in each plant. What is the average concentration in the third WT plant (genotype = WT, rep=3)?

Use the average concentration per plant for all subsequent analyses

2) Draw side-by-side boxplots of the concentrations in the three genotypes. What can you say about equal variances?

3) Log transform the average concentrations. What is the log transformed average for the third WT plant (genotype = WT, rep=3)?

4) Draw side-by-side boxplots of the log-transformed concentrations in the three genotypes. What can you say about equal variances?

Answers:

1) 0.3967 (more figures than appropriate so you know whether you’ve done the correct calculations)

2) It seems the HT genotype is more variable than the other two:



3) -0.924

4) The HT genotype still seems more variable, although somewhat less so than in the raw data.

