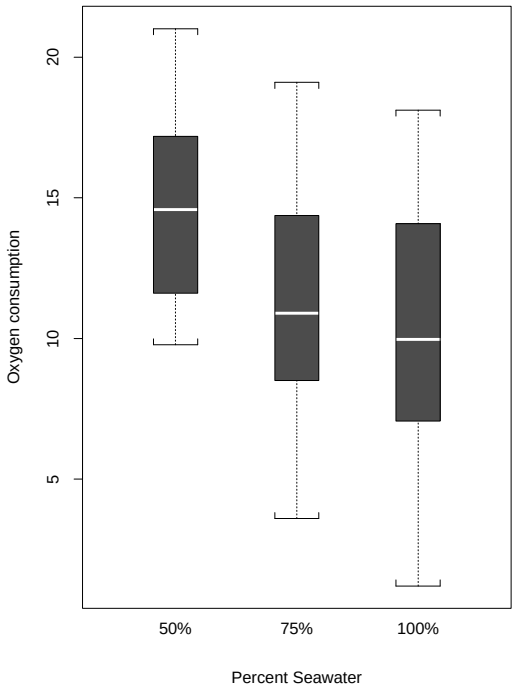
STA242/ENV255
Review of Prerequisite Concepts

Researchers were interested in whether seawater concentrations were related to oxygen consumption rates in limpets (molluscs that graze on algae).

A total of 48 limpets was used. Each of the limpets from each species was randomly assigned to one of the 3 concentrations of seawater (50%, 75% and 100%). Thus there are 16 limpets per seaweed level. The oxygen consumption rate was then measured in units of microliter oxygen per mg dry body weight per minute at 22 degrees C.



Partial Splus output follows:

<u>Analysis A</u>: Use a t−test to examine the difference between the 50% seaweed group and the 75% seaweed group.

Assume for this analysis that the data for the 100% group are not available.

Sample size for 50% group: 16
Sample size for 75% group: 16
Mean of 50% group: 14.9
Mean of 75% group: 11.3

standard deviation of 50% group: 3.9
standard deviation of 75% group: 4.3

T−test Splus output:

```
       Standard Two-Sample t-Test

data:  gp50 and gp75
t = ?, df = ?, p-value = ?
alternative hypothesis: true difference in means is not equal to
0
95 percent confidence interval:
 ?  ?
sample estimates:
 mean of x mean of y
   14.9       11.3
```

Analysis B: Use an ANOVA to examine how oxygen consumption differs among the three levels of seaweed.

Now assume that data for all three groups are available.

ANOVA:

```
             Df Sum of Sq  Mean Sq  F Value    Pr(F)
seawater      ?  187.6469  93.8235      ?       ?
Residuals     ?  798.8293  17.7518      ?
```

Review Questions:

1. Using the boxplot of the 50% seawater group, roughly estimate the median, interquartile range, upper quartile, lower quartile, minimum and maximum of the data. Compare the variability in the three groups. (Ch 1 of *Sleuth,* Section 1.2 of *M&M)*

2. Give the research question for Analysis A and Analysis B. For Analysis A, would it be appropriate to say that you are using a t−test to compare the difference in sample means? (Read pgs 256−257 of *M&M*)

3. Experimental design: Use Display 1.5 and associated discussion in *Sleuth*. Is enough information given to conclude that this a random sample? If not, how could you ensure that you have a random sample? Is this a randomized study?

4. Assumptions fall into 3 categories: Independence, Normality, and Variance. Give the assumptions of Analysis A and Analysis B. *State these in terms of the problem (independence of what exactly? normality of what exactly? variance of what exactly?).* Are you looking for normality of sample values or population values? Values of what? (Ch. 2 of *Sleuth)*

5. Give a way in which each of the three assumptions above could be violated for Analysis A and Analysis B. *State your answer in terms of the problem.* (Section 3.2.4 of *Sleuth)*

6. Write out the hypotheses for Analysis A and Analysis B. Are you writing the hypotheses in terms of $\overline{X}$ or μ? Which of these two quantities are fixed and which are random? Give the number of parameters that will be estimated in each analysis. (Section 7.2 of *M&M* and Ch 12 of *Sleuth)*

7. Write out a model statement for Analysis B in terms of the unknown parameters. (Ch 12 of *M&M*)

8. For Analysis A and Analysis B, calculate the test statistic. What distribution will you compare the test statistic to (be specific, giving distribution and degrees of freedom)? When you fill out the ANOVA table for Analysis B, does the "seawater" term have 1 degree of freedom since there is only 1 seawater variable?

9. Using $\alpha = 0.05$, what do you conclude from your result in (8) for Analysis A and Analysis B? Put your conclusions in terms of the problem at hand.

10. For each of Analysis A and Analysis B, give a 95% confidence interval for the unknown difference in mean oxygen consumption between the 50% seawater solution and the 75% seawater solution. Carefully write out the meaning of the confidence interval in terms of the problem.

11. What is the scope of inference for Analysis A and B? Again refer to Display 1.5 in *Sleuth* as well as Section 2.7 in *M&M*.