ANCOVAs

For Homework #8, you used simple ANOVAs for this data – here get more advanced on the same data to figure it out better!

At first glance, herbivory may be expected to reduce success of plants. But some plants, like *Ipomopsis aggregata* (to the right) actually *increase* in abundance when grazed. How does this happen? Hypotheses include:

1. Grazers avoid *Ipomopsis* and prefer other plants (competitive release)
2. Grazers eat *Ipomopsis* and other plants to reduce fitness of all, but other plants recover slower
3. Relative fitness of grazed *Ipomopsis* is actually increased relative to ungrazed *Ipomopsis*, for example **by increasing fruit production** as a response (overcompensation).

An experiment (Bergelson, J, Crawley, M. 1992. *American Naturalist* 139:870-882) was conducted on 14 natural populations in 5 states of the NW USA to test Hypothesis 3, which has been debated because it seems counterintuitive. Plants were first measured (including rootstock diameter), and then growing tips of plants were either left alone (`Ungrazed`) or clipped with scissors (`Grazed`). Choice of treatment was random among plants in a population.

Treatments were either `Grazed` or `Ungrazed`, and responses were the number of fruits per plant (`Fruits`), and the measured diameter of rootstock per plant (`Root`).

We first plot and analyze results as a simple ANOVA, as if Bergelson & Crawley did not have the foresight to measure plant size (because bigger plants can make more fruits).

Then we analyze again as a ANCOVA, accounting for differences in plant size. We compare models with AIC and by model selection.

1. Import and attach the `Ipomopsis` data set from the course web site.

2. Check assumptions of normality and homoscedasticity on Grazing subsets for fruit.

3. If assumptions are cool, compute a simple one-way ANOVA like you did in Homework #8 (call it `model1`), where `Fruit` is the response variable and `Grazing` is the treatment.

Is Hypothesis #3 supported so far?

But what is the relationship between `Root` and `Grazing`? Is this a problem? Do you expect bigger plants to produce more fruits regardless of grazing?

4. Try this plot to visualize the effect of `Root` on the results:

```
plot(Root,Fruit,pch=16+as.numeric(Grazing), col=c("red","black")
```

`[as.numeric(Grazing)])`

where red triangles = grazed plants and black diamonds = ungrazed plants

For a given root size (e.g., 7), are grazed plants actually producing more fruits?
And this looks like two *different* regressions, right? Your answer to the question ("For a given root size...") and the separation of symbols shows the effect of this covariate – the quantitative variable Root *covaries* with the results and needs to be included in the model.

Root might interact with Grazing or it could be a simple additive effect. Here we evaluate that choice using AIC.

BUT WAIT! How do we represent a combination of quantitative (e.g, Root) and categorical (e.g., Grazing) predictors? Which comes first? **The order that variables are listed matters for ANCOVA! We list the covariate first, categorical treatment second.**

Think of it this way: we compute an ANOVA *after* removing the signal due to the covariate. This is NOT a problem if only quantitative predictors are used in multiple regression.

To test this problem try a model as `Fruit ~ Root + Grazing`
and another as `Fruit ~ Grazing + Root`

Do you get different answers?

AIC-based Analyses. Now let's see how the three models to compare.

4. Load the `bbmle` package and use `AICctab` as in prior labs to generate an AIC table showing which model is most plausible. Options =
   1. `Fruit ~ Grazing + Root`
   2. `Fruit ~ Grazing`
   3. `Fruit ~ Root`
5. Which model has the greatest `weight`? A rule of thumb – if w_i is at least 2 X then next w, then it is clearly the most plausible model.

Consider again the plot with red triangles and black diamonds; do the results make sense?

So in the end, was Hypothesis #3 supported?
See how useful ANCOVA can be to correctly test hypotheses?

NOTE: the Root variable had to be measured – that means Bergelson & Crawley had to think of it ***before*** *they first collected data.* Lots of thought in advance, coupled with preliminary sampling and measurements led to a much better understanding of how plants respond to grazing. Our analyses reveal the effects for the well-designed study, and show how confusion over Hypothesis 3 was really confusion over how to conduct an experiment.