Quiz 7

1 – Consider the following example. You have 9 fields. Each field has 3 plots and each plot is randomly assigned to 1 of 3 treatments, in such a way that each field receives all treatments (i.e., randomized block design). What is the consequence of not including a random blocking variable for field in the statistical model (i.e., your analysis)?

* No consequence
* Pseudoreplication
* **Lower power**
* Autocorrelation

**It can be helpful to draw these out. 9 fields, each with 3 plots, each plot w/ a randomly assigned treatment. You don’t have to include a blocking effect in the field, and you can just run ANOVA. But you will lose power. Putting field in the model explains residual variation, which gives you smaller confidence intervals and smaller p-values.**

2 - Consider a similar experiment in which you have 27 fields, each with a single experimental plot in it. Each plot receives one of three treatments: control, treatment 1, and treatment 2. Which of the following is the appropriate statistical model (i.e., y ~ b0 + …) for analyzing data from this design?

* Size ~ B0 + B1\*Treatment1 + B2\*Treatment2 + E\_field~N(0,s\_field) + E\_residual~N(0,s\_residual)
* Size ~ B0 + B1\*Treatment1 + B2\*Treatment2 + E\_treatment~N(0,s\_treatment) + E\_residual~N(0,s\_residual)
* Size ~ B0 + B1\*Treatment1 + B2\*Treatment2 + E\_field\*treatment~N(0,s\_field\*treatment) + E\_field~N(0,s\_field) + E\_residual~N(0,s\_residual)
* **Size ~ B0 + B1\*Treatment1 + B2\*Treatment2 + E\_residual~N(0,s)**

**Again, we can draw the fields, each field only gets one plot with one treatment. The correct model only has a residual error. People get confused and want to put a field error term into here, but why can’t we? Since each field was only measured once and we only have 1 observation per field, we have nothing with which to measure variation due to field, so we can’t include it as a random effect. The correct answer is the last one: just a good-old fashioned ANOVA design.**

3 – Consider a study in which we measured the size of 32 individual trees each year from age 10 to age 20. What should be the random variable in your analysis?

* Size
* Age
* **Tree**
* Year

**A lot of people want to put ‘Age’ as the answer, but Age is a fixed effect. We want to know what the effect of Age is! Tree is the random effect, we have measured each tree multiple times, and we want to account for variation in size because some trees are just bigger than others.**

4 – Continuing from previous question, write an ‘lme()’ command in R that assumes the age-size relationship is different for each tree.

* **results <- lme(Size ~ Age, data = datum, random = ~Age|Tree)**
* results <- lme(Size ~ Age, data = datum, random = ~1|Age)
* results <- lme(Size ~ Age, data = datum, random = ~1|Tree)
* results <- lm(Size ~ Age, data = datum)

**In this case, we would want to have an interaction between the effect of Age and Tree. The appropriate model has a random effect of age, where each tree can have a different age-size relationship. Compared to the 1|TREE model, where each tree has a different Y-intercept for size, but it assumes they all have the same age-size relationship.**

5 – Continuing from the last two questions… True or False: there is no need to worry about autocorrelation in this system as autocorrelation is not possible.

* True
* **False**

**False! Because we have measured the tree repeatedly over time, there is definitely a possibility for autocorrelation.**