Brain dump

<https://highstat.com/index.php/mixed-effects-models-and-extensions-in-ecology-with-r>

(this is what’s in my markdown - Zoe)

Largely adapted from Kyle Edward's Lectures 21-22 and Ch. 13 of the Zurr (2009) book.

Luckily, most of the footwork for GLMMs has already been done for us, thanks to Tatum and Ana (GLMs) + Robert and Mallory (LMMs). We will be putting the two concepts together and showing you some fun and informative ways to visualize model predictions!

Robert and Mallory walked us very nicely through when we would need to use a mixed model structure, incorporating both random and fixed effects, but as a reminder, we incorporate random effects in the following situations:

1. Nested/hierarchical model structure

2. Inherently clustered data

3. Repeated measurements

Unfortunately for ecologists, our data is usually complicated beyond even the above. Assuming a Gaussian distribution simply doesn't work for most of the types of data we collect.

Is my sampling design nested, clustered, or repeated?

What <i>kind</i> of data do I collect? Count data? Presence/absense? Proportions/percents? Data with non-normal variance/error structure?

If the answer to both is YES! then you need generalized linear mixed models.

Instead of repeating all of the lovely work done by our predecessors, Sam and I will be covering some unique problems encountered when running GLMMs.

Issue 1: Many different ways to estimate parameters (both random and fixed) beyond simple maximum likelihood

Best options are Bayesian (Monte Carlo Maximum Likelihood) but this often deters ecologists.

Issue 2: Inference (most important to ecologists): WHICH MODEL IS BEST? WHAT DOES IT ALL MEAN?

- Wald Z test, chi squared - for normal variance

- Wald t test, F test - for overdispersion