ANCOVAs

ANCOVAs are hybrids of regression and ANOVA; a mix of categorical and continuous predictors for a response variable. ANCOVAs assume the effect of the continuous variable is constant across the categories (i.e., slopes are parallel). Your helicopter experiment mixed:

- categorical treatments = folded / not folded, body width (min & max)
- continuous treatment = wing length

We also had groups (a block effect) and a covariate (steps). Those are different, and represent a planned, and unplanned structure to the data collection, respectively.

You have already analyzed these data, but today we pay particular attention to two important details:

- the order in which you list categorical and continuous factors matters **for aov (not lm)** with a mix of quantitative and categorical predictors and where a study is unbalanced in replicates. This happens because R assumes sequentially-calculated ANOVAs (Type I models).
- ANCOVA assumes the continuous factor does not interact with other treatments (i.e., its slope is the same across other treatments)

1. Import and attach the `copters` data set from the course web site.
   https://sciences.ucf.edu/biology/d4lab/wp-content/uploads/sites/125/2019/09/helicopter-data.txt
   I assume you name the imported file to be "data" below.

2. Don't sweat assumptions today.

4. To remember / investigate pattern first, plot effect of WL on the results:

```
require(ggplot2)
ggplot(data, aes(x=WL, y=Time, color=Fold)) +
   geom_point(size=3) +
   geom_smooth(method="lm")
```

Does this look like it could be two *different* regressions?

5. And are the two regressions about parallel; *a requirement for ANCOVA?* How would you find out if slopes are equivalent between Folded and Not Folded? Go for it!

If slopes are parallel, then you have a simple ANCOVA. But if slopes are not parallel, then something else is going on, too, like an interaction, and we should probably not call this ANCOVA. And we know the experiment was more complicated. We'll get to that. But first,

If you are going to include both quantitative and categorical predictors in a model, which comes first? **The order that variables are listed is important for aov, so list the covariate first, categorical treatment second.** Think of it this way: we level the playing field for a fair comparison among categories by first removing the signal due to the quantitative "covariate."

5. To test this problem try one **aov** model as `Time ~ WL + Fold`
   and another as `Time ~ Fold + WL`

   Do you get different answers? What if you use **lm** instead?

AIC-based Analyses. Now let's see how the regression / ANOVA / ANCOVA compare.

6. Load the `bbmle` package and use `AICctab` as in prior labs to generate an AICc table showing which model is most plausible. Options =

```
a) Time ~ WL + Fold
b) Time ~ Fold + WL
c) Time ~ WL
d) Time ~ Fold
e) Time ~ 1
```

Which model has the greatest `weight`? What does this tell you?

Now let's consider how we should treat WL: as a category or a continuous covariate? First run a model with WL as a `cov`ariate (i.e., not interacting with other factors; ditto for Step):

```
covmodel <- lm(Time ~ WL + Step + Fold*BW + GROUP, data)
```

Now try it with WL as a `fac`tor that interacts with other factors:

```
fwing <- factor(data$WL)
facmodel <- lm(Time ~ Step + fwing*Fold*BW + GROUP, data)
```

And compare those by AICc:

```
AICctab(covmodel, facmodel, base=T, delta=T, weights = T)
```

This tells us that treating the interactions between WL and other treatments was important. So we shouldn't call this ANCOVA because WL interacted with other treatments (i.e., WL was not a constant effect, with parallel slopes).

So do you like representing wing lengths better as a factor or as a `cont`inuous variable?

```
contmodel <- lm(Time ~ Step + WL*Fold*BW + GROUP, data)
AICctab(covmodel, facmodel, contmodel, base=T, delta=T, weights = T)
summary(contmodel)
summary(facmodel)
```