Using AICc

The Akaike Information Criteria (AIC) will be used the rest of the semester and is a key part of "the new statistics." The fundamental goal: find the model – among your list of alternatives – that is most plausible. Note that says nothing about other possible models that are not listed. AICs can be applied to categorical predictors (as used in ANOVAs), continuous predictors (as used in regression), or combinations of both.

We have already seen how models with > N and/or > variables can affect p values and variance "explained" as measured by R^2. Thus R^2 is helpful *but not a fair way to compare models*. AIC discounts models for the number of variables to find the most parsimonious or most efficient model. We won't worry about the math here, but need to understand the output. Reported metrics include:

- AIC or corrected AIC (AICc). The AICc should be your default, because it corrects for low N and equals AIC at large N. Lower values indicate more plausible models.
- delta AICc. The difference between ranked models. A delta AICc ~ 2 indicates a clear choice – otherwise, two models are comparable.
- AICc weight. This represents the relative likelihood of a model, where 1.0 = most likely. *Weight is the best way to rank and compare models.*

Multiple R packages report AIC metrics, including bbmle and AICcmodavg, which produce simple tables to compare models. Here we use bbmle because it is simple to code.

Load and attach our copter data from:
http://jenkins.cos.ucf.edu/wordpress/wp-content/uploads/copter-data-F16.csv

And make fold, wing, and group factors

```
ffold <- factor(Fold)
fwing <- factor(Wing)
fgroup <- factor(Group)
```

Consider different alternative models for our copter experiment:

```
nullmodel <- lm(Time ~ 1) # This says Time is a constant across all
   # treatments. Other models that are less plausible than the null obviously
don't tell us much
groupstepmodel <- lm(Time ~ Group + Step) # This says only groups & steps
mattered –
   # not really expected, but compares our "other" factors to planned factors
wingmodel <- lm(Time ~ fwing) # This says only wings mattered
foldmodel <- lm(Time ~ ffold) # This says only folds mattered
wingfoldmodel <- lm(Time ~ fwing*ffold) # Closer to our hypothesized model
fullmodel  <- lm(Time ~ fwing*ffold + fgroup + Step) # Our hypothesized
experimental design
```

Notice that not all possible model combinations are listed? Why? Because these are all I

hypothesized. **AIC only compares models you list** – you have to think first about which ones you are interested in, for good reasons.
So to be clear, we conducted the copters experiment using the full model and so we expected *a priori* that the full model would be most plausible (otherwise we would have conducted another experiment). But maybe we were wrong, and factors that do not help explain the variance may be left out for a simpler, more plausible model. Let's find out:

Now install (if not already done) and load the `bbmle` package.

Now run this command to generate a table for AICc scores, etc.

```
AICctab(nullmodel, groupstepmodel, wingmodel, foldmodel, wingfoldmodel,
fullmodel, base=T, delta=T, weights=T)
# This simply asks for a AICc table of the listed models, where the table will
include base AICc values, delta AICc value, and weights
```

Clearly the full model kicks butt (technically speaking) because it is at least 1000x more plausible than other models, despite being more complex. So we are justified in presenting the full ANOVA and lm results:

```
summary(fullmodel)
summary.aov(fullmodel)
```

Now let's use AICctab for another data set.

Get the data set:

1. In the MASS package, there is a data set on 1993 cars. If MASS is not already installed, install it now, and then load MASS to get access to Cars93. If MASS is already installed, then simply load it.
2. Because it comes with a package, we load Cars93 differently than if when we import a txt file:

```
data(Cars93)
attach(Cars93)
```

I start with two hypotheses and compare them with AICc – as a template to show you how to proceed. Then you make three more models and compare them all.

Bet 1: I bet MPG.city is simply predicted by Origin (US vs non-US cars). Thus my model 1 and request for output looks like this:

```
model1 <- lm(MPG.city ~ Origin)
summary.aov(model1)
boxplot(MPG.city ~ Origin)
```

Bet 2: I bet MPG.city in 1993 can be simply predicted by Manufacturer.

Make a set of statements similar to Bet 1 above to match.

Examine the Adjusted R^2 of the two models using `summary(modelX)`
Which model would you think is the best?
Now compare those two models using AICc instead:

```
AICctab(model1, model2, weights = TRUE, delta = TRUE, base=TRUE, sort = TRUE)
```

So... While both factors (Origin, Manufacturer) significantly affected MPG.city, the Origin-based more parsimoniously explains MPG.city. AIC-based model selection discounts a model for the number of terms in it – thus it assesses “bang for the buck” - the most *efficient* models are most plausible.

What other factors might also affect MPG.city? **Construct at least THREE MORE alternative models** to evaluate. Make models as complex as you think is required, BUT a model should represent a hypothesis – such as my bets above. *Grab-bag / smörgåsbord / all-possible-options models do not count because this approach is about testing* a priori *hypotheses*. Evaluate each model (as above), and then compare all the models using AICctab.

Also take note of the Adjusted R^2 of the models you evaluate (but do not use that as a criterion to select the most plausible model). And notice that it would be useful to report models' coefficients, adjusted R^2 values, and other output, but only *after* most-plausible models were selected by AICc.