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.

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 with different explanatory variables*. AIC discounts models for the number of variables to find the **most plausible model**. 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. 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 (w_i). This represents the relative likelihood of a model, where 1.0 = most likely. ***Weight is the best way to rank and compare models***.

Load and attach our copter data from:
https://sciences.ucf.edu/biology/d4lab/wp-content/uploads/sites/125/2019/09/helicopter-data.txt

And make fold, wing, and group factors for categorical treatments (ANOVA-style output):

```
ffold<- factor(Fold)
fwl <- factor(WL)
fbw <- factor(BW)
fgroup <- factor(GROUP)
```

The first, important step in using AICs is to carefully construct hypotheses to test. Consider these for our copter experiment:

```
nullmodel <- lm(Time ~ 1) # This says Time is a constant.
   # Always include nulls in your model sets!!!
groupstepmodel <- lm(Time ~ fgroup + Step) # “only groups & steps mattered”
wingmodel <- lm(Time ~ fwl + fgroup) # “only wings mattered”
foldmodel <- lm(Time ~ ffold + fgroup) # “only folds mattered”
wingfoldmodel <- lm(Time ~ fwl*fbw*ffold + fgroup) # hypothesized model
ignoring step
fullmodel  <- lm(Time ~ Step + fwl*fbw*ffold + fgroup) # Our full,
hypothesized experimental design that accounts for steps too
```

Notice that not all possible model combinations are listed? Why? Because these are all ***I*** hypothesized. You wanna make more hypotheses? Go ahead! But choose wisely: To throw all possible models at a question is like fishing with dynamite – you may get something, but you didn't do Science (i.e., test ideas) and you have to live with the consequences.

AIC only compares models you list – you have to think first about which ones you are interested in, for good reasons.

To be clear, the full model best represents the experiment as conducted, and so we expect *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.

To compute AICs, 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
# ACTUALLY list all the models - including ones you made too
# AIC tables in papers customarily list all these measures
```

***See the 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)
```

<u>**Now let's use AICctab for another data set.**</u>

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. If MASS is already installed, then turn it on (check the box).
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.

<u>Bet 1</u>: 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)
```

<u>Bet 2</u>: 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* **<u>a priori</u>** *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 you should report model coefficients, adjusted R^2 values, and other output for your most-plausible model(s) identified by AICc.