Which bag? – Likelihood and AIC

This is a quick game to introduce the concepts of likelihood and AIC.

The experiment

We have a collection of identical small bags each containing Go stones. Some of the bags contain only white stones; in others, 9 of the stones are white and 1 is black. Take a bag at random: don't look inside!

The aim is to decide whether the bag you took contains all white stones or mixed stones by pulling out one stone at random. Shake up the bag, put in your hand, and pull out one stone.

- If the first stone is black, you can be certain that you have a 'mixed' bag, as it is impossible to pull a black stone out of an 'all white' bag.
- But if the first stone is white, you can't be certain either way, though you have evidence that it is an 'all white' bag.

Do you think drawing 1 white stone is good evidence for an 'all white' bag? Or is the '1 black' hypothesis still plausible?

Quantifying the evidence

We'll use a spreadsheet (using Excel or Calc or other software) to calculate AIC for the 2 models: (a) all white, (b) 9 white, 1 black. If you drew a white stone the spreadsheet will look like this:

Draw:	1	
Colour:	w	
Likelihood		
all white	1	This is the probability of drawing a white stone for
9w, 1b	0.9	each of the 2 models.
AIC		
all white	0	We have 0 parameters to estimate, so AIC = - 2*LN(likelihood).
9w, 1b	0.21	
delta AIC	0.21	The all-white model has lowest AIC.

As you draw more stones, you add more columns to the spreadsheet. This is what it would look like if you drew 3 white stones in a row, replacing each stone and shaking the bag before drawing the next:

Draw:	1	2	3
Colour:	w	w	w
Likelihood			
all white	1	1	1
9w, 1b	0.9	0.81	0.729
AIC			
all white	0	0	0
9w, 1b	0.21	0.421	0.632
delta AIC	0.21	0.421	0.632

For this use = 0.9 ^ C1, then you can copy to the right.

The difference in AIC increases as you draw more white stones.

What if you drew many times? How would you assess the evidence for the all-white bag after drawing white stones 10, 20, 30, 50 times? Add columns to the spreadsheet for these numbers of draws.

The table below is based on Burnham and Anderson's (2002, p.70) rules of thumb for models with different ΔAIC values:

ΔAIC	**no. of white stones**	**Inference**
0 – 2	< 9	Don't know which is the better model
4 – 7	19 – 33	Evidence that the lower AIC model is better
> 10	> 48	We can ignore the model with higher AIC

What if you draw a black stone?

In this case you know for sure that the bag must be 9 white, 1 black. Let's see what happens to the AIC values.

Make a copy of the spreadsheet. Keep the first few columns and then put in a column for a black stone.

Draw:	1	2	3	4
Colour:	w	w	w	**b**
Likelihood				
all white	1	1	1	0
9w, 1b	0.9	0.81	0.729	0.0729
AIC				
all white	0	0	0	Inf
9w, 1b	0.21	0.421	0.632	5.237
delta AIC	0.21	0.421	0.632	Inf

LN(0) gives an error; in R this would be Inf.

Now 9w-1b has lowest AIC, and delta AIC is infinite!

Notice that the likelihood for the 9w-1b model goes down gradually as we draw more white stones, then drops suddenly when we pull out a black stone. This sudden drop corresponds to the moment that we are sure the 9w-1b model is correct! The actual value of the likelihood doesn't tell us anything about the evidence for the model. When we pull out a black stone, the likelihood for the all-white model drops to zero, which is even lower than the 9w-1w model; we need to know both likelihoods to decide between the models.

Summary and outlook

- If you have to select one model from the set of models, the one with the lowest AIC is best.
- There is always uncertainty when we are drawing conclusions from a sample: a different randomly-drawn sample could lead to different conclusions. So we can't rule out the models with higher AIC.
- If the difference in AICs, ΔAIC, is small, the model is still plausible. With a different sample, it might turn out to be the best. If ΔAIC is large, it's most unlikely to ever come out top.
- In later modules we'll see how to deal with uncertainty in model predictions using AIC weights.

Reference

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*, 2 edn. Springer-Verlag.