AKAIKE'S INFORMATION CRITERIA (AIC)

The general form for calculating AIC:

$$\textbf{AIC} = -2*\ln(\text{likelihood}) + 2*K$$

where **ln** is the natural logarithm

(likelihood) is the value of the likelihood

K is the number of parameters in the model, e.g., consider the regression equation

$$\underset{1}{Growth = \underset{\wedge}{10}} + \underset{1}{\underset{\wedge}{5}}*age + \underset{1}{\underset{\wedge}{3}}*food + \underset{1}{\underset{\wedge}{error}}$$

$1 + 1 + 1 + 1 = 4$ parameters

AIC can also be calculated using residual sums of squares from regression:

$$\textbf{AIC} = n*\ln(RSS/n) + 2*K$$

where **n** is the number of data points (observations)

RSS is the residual sums of squares

AIC requires a bias-adjustment small sample sizes. B&A rule of thumb: If ratio of n/K < 40, then use bias adjustment:

$$AIC_C = -2*\ln(\text{likelihood}) + 2*K + (2*K*(K+1))/(n-K-1)$$

where variables are as defined above. Notice that as the size of the dataset, n, increases relative to the number of parameters, **K**, the bias adjustment term on the right becomes very, very small. Therefore, it is recommended that you always use the small sample adjustment.

For example, consider 3 *candidate models* for the growth model above, their RSS values, and assume n = 100 samples in the data:

Model	K	RSS	AIC_c
Food, Age	**4**	*25*	100*ln(*25*/100) + 2***4** + (2***4***(**4** + 1))/(100 **-** **4** -1) = -130.21
Food	**3**	*26*	100*ln(*26*/100) + 2***3** + (2***3***(**3** + 1))/(100- **3**- 1) = -128.46
Age	**3**	*27*	100*ln(*27*/100) + 2***3** + (2***3***(**3** + 1))/(100- **3**- 1) = -124.68

MODEL SELECTION WITH AIC

The best model is determined by examining their relative distance to the "truth". The first step is to calculate the difference between model with the lowest AIC and the others as:

$$\boldsymbol{\Delta_i = AIC_i - \min AIC}$$

where Δ_i is the difference between the AIC of the best fitting model and that of model *i*

AIC_i is AIC for model *i*

min AIC is the minimum AIC value of all models

For example, consider the 3 *candidate models* and their AIC_c values:

Model	K	RSS	AIC_c
Food, Age	4	25	-130.21
Food	3	26	-128.46
Age	3	27	-124.68

The smallest value is for the model containing *Age* and *Food* with –130.21. Thus the Δ_i are:

Model	K	RSS	AIC_c	Δ_i
Food, Age	4	25	-130.21	-130.21 + 130.21 = 0.00
Food	3	26	-114.15	-128.46 + 130.21 = 1.75
Age	3	27	-98.73	-124.68 + 130.21 = 5.52

(Note 130.21 is added because subtracting a negative number = addition.)

For publication purposes, candidate models are always arranged in ascending order of Δ_i as is shown above.

To quantify the plausibility of each model as being the best approximating, we need an estimate of the likelihood of our model given our data.

$$\boldsymbol{\mathcal{L}(\text{model}|\ \text{data})}$$

Interestingly, this proportional ($\propto$) to the exponent of $-0.5*\Delta_i$ so that

$$\boldsymbol{\mathcal{L}(\text{model}|\ \text{data}) \propto \exp(-0.5*\Delta_i)}$$

The right hand side of above is known as the *relative likelihood* of the model, given the data.

MODEL SELECTION WITH AIC (CONT)

A better means of interpreting the data is to normalize the relative likelihood values as:

$$w_i = \frac{exp(-0.5 * \Delta_i)}{\sum_{r=1}^{R} exp(-0.5 * \Delta_r)}$$

where w_i are known as *Akaike weights* for model *I* and the denominator is simply the sum of the relative likelihoods for all candidate models.

For example, using the earlier values from the 3 growth models:

Model	K	RSS	AIC_c	Δ_i	$exp(-0.5*\Delta_i)$
Food, Age	4	25	-130.21	0	1.0000
Food	3	26	-128.46	1.75	0.4166
Age	3	27	-124.68	5.52	0.0631
				Sum =	1.4798

The sum of the relative likelihoods is 1.4798, so we obtain the Akaike weights for each by dividing the relative likelihood by 1.4798.

Model	K	RSS	AIC_c	Δ_i	w_i
Food, Age	4	25	-130.21	0	0.6758
Food	3	26	-128.46	1.75	0.2816
Age	3	27	-124.68	5.52	0.0427

The above example table is the recommended format for publication. We now interpret the w_i as the weight of evidence that model *i* is the best approximating model, given the data and set of candidate models. Alternatively, the w_i can be interpreted as the probability that *i* is the best model, given the data and set of candidate models. For the above example, the model containing age and food is (0.6758/0.2816) = 2.4 times more likely to be the best explanation for growth compared to food only and (0.6758/0.0427) = 15.8 times more likely than age only. As a general rule of thumb, the *confidence set* of candidate models (analogous to a confidence interval for a mean estimate) include models with Akaike weights that are within 10% of the highest, which is comparable with the minimum cutoff point (i.e., 8 or 1/8) suggested by Royall (1997) as a general rule-of-thumb for evaluating strength of evidence. For the above example, this would include any candidate model with a value greater than (0.6758*0.10) = 0.0676. Thus, we would probably exclude the model containing age only from the model confidence set because its weight, 0.0427<0.0676. The conclusion would be that there was insufficient evidence to consider age only as a plausible explanation for growth.

AKAIKE IMPORTANCE WEIGHTS FOR PARAMETERS

The relative importance of individual parameters can also be examined using Akaike weights. Here, the Akaike weights for each model that contains the parameter of interest are summed. For the growth models (above), the importance weights would be:

	Candidate model			
Parameter	Food and Age	Food only	Age only	Importance weight
Food	0.6758	+ 0.2816	+ 0.0000	= 0.9573
Age	0.6758	+ 0.0000	+ 0.0427	= 0.7184

Food and age are both highly plausible explanations for growth. However, food is (0.9573/0.7184) = 1.33 times more plausible, given the data and candidate models.

MODEL SELECTION UNCERTAINTY AND PARAMETER ESTIMATES

Often the parameter estimates (e.g., slope and intercepts in regression models) for the same variable in differ among candidate models. For example,

Age & Food model

$$Growth = 10 + 3*food + 5*age + error$$

Food only model

$$Growth = 15 + 7*food + error$$

Age only model

$$Growth = 12 + 10*age + error$$

Notice that the parameter estimate for *food* is 3 and 7 for the “age and food” and “food only” models, respectively, and that of *age* is 5 and 10 for the “age and food” and “age only” models, respectively.

MODEL SELECTION UNCERTAINTY AND PARAMETER ESTIMATES (CONT)

Which estimate of the effect of *food* and *age* on growth is correct? Maybe we should just pick the values from the most plausible model, the "food and age" model. However, the Akaike weight for the "food only" model (0.282) tells us that this model is still a plausible explanation for growth, given the data and set of candidate models.

What about simply averaging the values of the models?

Why would we want to give equal weight to each model when we know some are better than others?

The idea behind AIC model averaging is to use the Akaike weights to *weight* the parameter estimates and variances (i.e., standard errors) from each model and combine those. Thus, we incorporate model selection uncertainty directly into the parameter estimates via the Akaike weights.

Model-averaged parameter estimates are only calculated for those parameters (variables) that are included in the confidence set of models. For the growth example, the *intercept*, *food*, and *age* are contained in the model confidence set, that is, they're in the "food and age" and "food only" models. There are two methods for model-averaging- $\bar{\hat{\beta}}_j$, where parameter estimates are averaged over all models in which predictor x_j occurs and $\bar{\tilde{\beta}}_j$, where parameter estimates are averaged over all models not just those in which predictor x_j occurs.

Model averaged parameter estimates under $\bar{\hat{\beta}}_j$ are calculated in 4 simple steps.

Step 1: Use the exponentiated AIC values, exp(-0.5*Δ_i), only from the models that contain the parameter.

Step 2: Akaike weights need to sum to 1 (just like a probability), so add the exp(-0.5*Δ_i) values from all of the candidate models containing the parameter to get a new sum.

Step 3: Divide the exp(-0.5*Δ_i) by new sum to get new Akaike weights.

Step 4: Multiply the raw (individual model) parameter estimates by the new weights and sum.

MODEL SELECTION UNCERTAINTY AND PARAMETER ESTIMATES (CONT)

These steps applied to the growth model are illustrated below with model-averaged estimates shown in bold.

Model	$\exp(-0.5*\Delta_i)$	New weight ($\exp(-0.5*\Delta_i)$/sum)		Raw parameter estimate		Weighted parameter estimate
Intercept estimate						
Age, Food	1.0000	0.6758	*	10	=	6.758
Food	0.4166	0.2815	*	15	=	4.223
Age	0.0631	0.0426	*	12	=	0.512
sum =	1.4798				sum =	**11.492**
Food estimate						
Age, Food	1.0000	0.7059	*	3	=	2.118
Food	0.4166	0.2941	*	7	=	2.059
sum =	1.4166				sum =	**4.176**
Age estimate						
Age, Food	1.0000	0.9406	*	5	=	4.703
Age	0.0631	0.0594	*	10	=	0.594
sum =	1.0631				sum =	**5.297**

Thus we have the *composite model* for growth

$$Growth = \mathbf{11.492} + \boldsymbol{4.176}*food + \mathbf{5.297}*age + error$$

MODEL SELECTION UNCERTAINTY AND PARAMETER ESTIMATES (CONT)

Parameter estimates are also estimated with a certain amount of error that, in computer outputs, is reported as the *standard error* of the estimate. The standard error is important because it is used to determine the reliability of the parameter estimate. Large standard errors (generally, 2X > parameter estimate) mean that the parameter estimate is not reliable for predicting the outcome or interpreting the model. Below are the outputs for each of the candidate models of growth.

Food and age model

Parameter	Estimate	Standard Error
Intercept	10.000	2.000
Food	3.000	0.500
Age	5.000	2.500

Food only model

Parameter	Estimate	Standard Error
Intercept	15.000	5.000
Food	7.000	1.500

Age only model

Parameter	Estimate	Standard Error
Intercept	12.000	3.000
Age	10.000	1.500

Model-averaged parameter estimates should always have a measure of reliability. These are calculated similar to the model-average parameter estimates in that the used Akaike weights to weight the standard errors from each candidate model (above). However, these standard errors are *conditional* on the candidate model. Therefore, an additional source of variance, the *model selection variance,* must be included.

MODEL SELECTION UNCERTAINTY AND PARAMETER ESTIMATES (CONT)

Model selection variance (MSV) is estimated using the model-averaged estimate and the raw parameter estimates from the candidate models and is calculated as:

$$\text{MSV} = (\text{model-averaged estimate} - \text{raw parameter estimate})^2$$

Estimates of model selection variance for the growth model are illustrated below.

Model	Model-averaged estimate		Raw parameter estimate		Model selection variance
Intercept estimate					
Age, Food	(11.492	–	$10)^2$	=	2.227
Food	(11.492	–	$15)^2$	=	12.304
Age	(11.492	–	$12)^2$	=	0.258
Food estimate					
Age, Food	(4.176	–	$3)^2$	=	1.384
Food	(4.176	–	$7)^2$	=	7.973
Age estimate					
Age, Food	(5.297	–	$5)^2$	=	0.088
Age	(5.297	–	$10)^2$	=	22.120

To calculate the unconditional standard errors, the model selection variance is added to the conditional variance (the model standard errors squared). The square root of this sum is then weighted by the Akaike weights and summed, similar to the model average parameter estimates.

MODEL SELECTION UNCERTAINTY AND PARAMETER ESTIMATES (CONT)

These steps applied to the growth model are illustrated below with model-averaged unconditional standard errors shown in bold.

Model	Standard Error	Conditional variance (Standard error)^2	Model selection variance	Square root of (Cond Var + MSV)		New weight	Weighted unconditional standard error
Intercept estimate							
Age, Food	2.000	4.000	2.227	2.495	*	0.6758	= 1.686
Food	0.500	0.250	12.304	3.543	*	0.2815	= 0.998
Age	2.500	6.250	0.258	2.551	*	0.0426	= 0.109
						sum =	**2.793**
Food estimate							
Age, Food	0.5	0.250	1.384	1.278	*	0.7059	= 0.902
Food	1.5	2.250	7.973	3.197	*	0.2941	= 0.940
						sum =	**1.842**
Age estimate							
Age, Food	2.5	6.250	0.088	2.518	*	0.9406	= 2.368
Age	1.5	2.250	22.12	4.937	*	0.0594	= 0.293
						sum =	**2.661**

For interpretation, the reliability (precision) of model averaged parameter estimates (MAE) should be reported with the aid of confidence intervals (CI) using the unconditional standard errors (SE). This can easily be accomplished as:

Upper CI = MAE + (*t-value**SE)
Lower CI = MAE - (*t-value**SE)

where *t-value* is the critical value from a t-distribution based on sample size and the confidence interval desired, e.g., the *t-value* for a 95% CI with 20 or more samples = 1.95 and the value for 90% CI with 20 or more samples = 1.64.

The below example table is recommended as the format for reporting the composite model in a publication.

			90% CI	
Parameter	Estimate	SE	Upper	Lower
Intercept	11.492	2.793	16.073	6.911
Food	4.176	1.842	7.197	1.155
Age	5.297	2.661	9.661	0.933

USEFUL REFERENCES

Burnham, K.P., and Anderson, D.R. 2002. Model selection and inference: a practical information-theoretic approach, second edition. Springer-Verlag, New York.

Royall, R.M. 1997. Statistical evidence: a likelihood paradigm. Chapman and Hall, New York.