

AICcTab Function

```
AICcTab(models,nobs=10,mnames=c("Null","Spatial","Seven","Spatial*Seven","Global"),base=T,weight
s=T,delta=T,logLik=T)
```

- Function needs
 - List of models
 - * List_of_models<-list(model1,model2,model3,model4...)
 - nobs – number of observations
 - mnames (options) list of model names for the output
 - weight (T/F) – include model weights
 - delta (T/F) – include delta AIC
 - logLik (T/F) – include log likelihood

Models

- For AIC analyses, models can be just about anything
- Often, these are simple linear models (lm function)
 - * First_model<-lm(response_var~predictor1+predictor2)
- For this assignment, your predictor variable is log_resratio
 - Already log transformed
 - Represents each species ratio of stream to reservoir abundance
 - Your assignment is to use AIC to select a parsimonious model predicting this from species life history traits.



Log_resratio=2.56



Log_resratio=0.0002

Specifying Sample Size in AICcTab

```
AICcTab(models,nobs=10,mnames=c("Null","Spatial","Seven","Spatial*Seven","Global"),base=T,weight
s=T,delta=T,logLik=T)
```

	logLik	AICc	dLogLik	dAICc	df	weight
Global	10.7	-67.1	67.7	0.0	16	1
Seven	2.2	5.6	59.3	72.7	3	<0.001
Spatial*Seven	7.6	54.8	64.6	121.9	7	<0.001
Spatial	-25.7	67.5	31.3	134.6	4	<0.001
Null	-57.0	119.8	0.0	186.9	2	<0.001

```
AICcTab(models,nobs=29,mnames=c("Null","Spatial","Seven","Spatial*Seven","Global"),base=T,weight
s=T,delta=T,logLik=T)
```

	logLik	AICc	dLogLik	dAICc	df	weight
Seven	2.2	2.5	59.3	0.0	3	0.69
Spatial*Seven	7.6	4.1	64.6	1.6	7	0.31
Global	10.7	55.9	67.7	53.4	16	<0.001
Spatial	-25.7	61.1	31.3	58.6	4	<0.001
Null	-57.0	118.5	0.0	116.0	2	<0.001

Not the same in AICtab function

```
AICtab(models,nobs=10,mnames=c("Null","Spatial","Seven","Spatial*Seven","Global"),base=T,weights
=T,delta=T,logLik=T)
```

	logLik	AIC	dLogLik	dAIC	df	weight
Spatial*Seven	7.6	-1.2	64.6	0.0	7	0.7984
Seven	2.2	1.6	59.3	2.8	3	0.1994
Global	10.7	10.6	67.7	11.8	16	0.0022
Spatial	-25.7	59.5	31.3	60.7	4	<0.001
Null	-57.0	118.1	0.0	119.3	2	<0.001

```
AICtab(models,nobs=29,mnames=c("Null","Spatial","Seven","Spatial*Seven","Global"),base=T,weights
=T,delta=T,logLik=T)
```

	logLik	AIC	dLogLik	dAIC	df	weight
Spatial*Seven	7.6	-1.2	64.6	0.0	7	0.7984
Seven	2.2	1.6	59.3	2.8	3	0.1994
Global	10.7	10.6	67.7	11.8	16	0.0022
Spatial	-25.7	59.5	31.3	60.7	4	<0.001
Null	-57.0	118.1	0.0	119.3	2	<0.001

If you are using linear models (lm), you may not need to

```
nobs (integer) number of observations: required for type="BIC" or type="AICc" unless objects
have a nobs method

AICcTab(models, nobs=29, mnames=c("Null", "Spatial", "Seven", "Spatial*Seven", "Global"), base=T, weight
s=T, delta=T, logLik=T)

      logLik AICc  dLogLik dAICc df weight
Seven      2.2  2.5  59.3    0.0  3  0.69
Spatial*Seven  7.6  4.1  64.6    1.6  7  0.31
Global     10.7  55.9  67.7   53.4 16 <0.001
Spatial    -25.7  61.1  31.3   58.6  4 <0.001
Null      -57.0 118.5   0.0  116.0  2 <0.001

> nobs(spatial)
[1] 29

AICcTab(models, mnames=c("Null", "Spatial", "Seven", "Spatial*Seven", "Global"), base=T, weights=T, del
ta=T, logLik=T)

      logLik AICc  dLogLik dAICc df weight
Seven      2.2  2.5  59.3    0.0  3  0.69
Spatial*Seven  7.6  4.1  64.6    1.6  7  0.31
Global     10.7  55.9  67.7   53.4 16 <0.001
Spatial    -25.7  61.1  31.3   58.6  4 <0.001
Null      -57.0 118.5   0.0  116.0  2 <0.001
```

Regression Tree Object

```
node), split, n, deviance, yval
* denotes terminal node

1) root 29 86.7551000 1.531342e-17
2) env7< 0.4530161 14 9.0540510 -1.565784e+00
4) env7<= 0.2482159 6 0.8007901 -2.356327e+00 *
5) env7>= 0.2482159 8 1.6911960 -9.728763e-01 *
3) env7>= 0.4530161 15 11.3422800 1.461398e+00
6) env7< 0.6909247 7 1.3414610 6.710477e-01 *
7) env7>= 0.6909247 8 1.8022420 2.152955e+00 *

Vector assigning samples to groups
> regtree$where
[1] 3 3 3 3 3 3 4 4 4 4 4 4 4 4 6 6 6 6 6 6 7 7 7 7 7 7 7 7

> regtree$plits
      count  best  improve  index adj
env7    29  -1  0.7648976 0.4530161  0
spatial 29   3  0.6793786 1.0000000  0
env2    29  -1  0.1877116 0.2115776  0
env5    29  -1  0.1630396 0.4046268  0
env6    29  -1  0.1499714 0.6479728  0
env7    14  -1  0.7287656 0.2482159  0
spatial 14   3  0.6649992 2.0000000  0
env1    14  -1  0.2698882 0.6441778  0
env4    14  -1  0.2570974 0.8397628  0
env2    14  -1  0.2560837 0.9360151  0
env7    15  -1  0.7228333 0.6909247  0
spatial 15   3  0.6710280 0.0000000  0
env3    15  -1  0.2238585 0.7973027  0
env5    15  -1  0.2072037 0.3884382  0
env1    15  -1  0.1680820 0.7431955  0
```

List of the candidate splits.
 Improve = decrease in deviance
 Index = where the split is made

Fine Tuning

```
rpart(response~var1+var2+var3, data=envdata, control=rpart.control(minsplit=15, minbucket=10))
```

- minsplit – minimum number of samples in a split (5)
- minbucket – minimum samples in a terminal bucket (2)
- surrogatestyle – how to choose splits, penalties for missing data or missclassifications
- cp – complexity parameter. Degree or required improvement in group homogeneity (0.01)
 - With lower cp, you will grow a larger tree that can be pruned with cross validation

- Control of boosting
 - Learning rate (shrinkage) – contribution of each subsequent tree to overall model
 - Complexity – number of nodes and interactions
 - Number of trees

Journal of Animal Ecology

Journal of Animal Ecology, 2002, 71, 851–857

A working guide to boosted regression trees

J. Elmer*, J. R. Leathwick* and T. Hayes†

*School of Botany, The University of Melbourne, Parkville, Victoria, Australia 3101; †National Institute of Water and Atmospheric Research, PO Box 11710, Hamilton, New Zealand; and ‡Department of Statistics, Stanford University, CA, USA

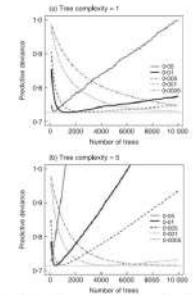


Fig. 2. The relationship between number of trees and predictive deviance for models fitted with the learning rates and two levels of tree complexity. Deviance was calculated from models fitted to a data set of 1000 observations, and predicted to a data set of 12 500 observations. The lowest predictive deviance achieved for each model is indicated by a dotted horizontal line; the line for learning rate achieving that minimum is shown in bold.

