AIC and all that

Ben Bolker

23:50 05 July 2015

Background

- novel approach to *model selection*
- $\underbrace{-2L}_{\text{badness of fit}} + \underbrace{2k}_{\text{complexity penalty}}$
- minimize expected difference
- asymptotically equivalent to *leave-one-out cross-validation*2
- typical method: fit all (?) possible models ("dredge")
- select best model (?)
- compute *AIC weights*
- multi-model averaging

Problem

McGill (2015), "Why AIC Appeals to Ecologist's Lowest Instincts": some goals of statistical ecological modeling are

> . . . estimation of parameters, testing of hypotheses, exploration of covariation, prediction into new conditions, selecting among choices (e.g. models) . . .
>
> (*could add "quantifying variable importance"*);
>
> With AIC you present that classic table of ΔAIC and weights and voila! You've sort of implied doing all five statistical goals at once.

Advantages

- simple
- handles non-nested models (but see Ripley (2004))
- accounts for model complexity

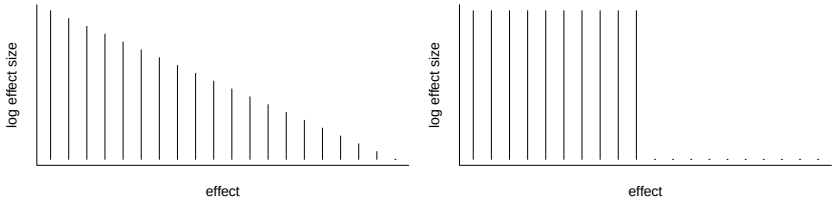
AIC goals

- AIC is for *prediction*
- is it OK to use it to test hypotheses?

ICs

- distinction among ICs
  - AIC

  - DIC: (Spiegelhalter et al. 2002): prediction for Bayesian multilevel models
  - BIC: identify true model (Yang 2005)
  - AICc: finite-sample correction (Richards 2005; Fletcher and Dillingham 2011)
  - QAIC: correct for overdispersion
  - CAIC (conditional) (also DIC): multilevel predictions (Vaida and Blanchard 2005; O'Hara 2007)
- statistical inconsistency/overfitting argument; consistency vs efficiency (Yang 2005) and *tapering effect sizes*: blog post



- still need to respect limits of model complexity (Harrell 2001)
- model-averaged CIs are a good idea, but still represent hypothesis testing

Model selection

- OK, but why?
- don't take $\Delta\text{AIC} > 2$ criterion too seriously

Multimodel averaging

- averaging *predictions* is completely OK
- parameter averaging (Cade 2015)
  - must average *predictions*
  - parameters in linear models *may* represent predictions
  - problems with multicollinearity; interaction terms (Schielzeth 2010); nonlinear models
  - how to average zero values?
- `MuMIn`, `AICcmodavg` packages
- multimodel averaging shrinkage vs. penalized regression (Lasso/elastic net/et al.)

Weights and variable importance

- are model weights probabilities? of what?

- prob. of inclusion in true best model?
- 'savvy priors'
- what would Bayesian probabilities mean?

Alternatives

- penalized regression not without challenge either, but much faster
- expanded models (with shrinkage?), i.e. don't test point hypotheses

Good practice

- consider adequacy
- present ΔAIC, not AIC (`bbmle::AICtab`)
- graphical results: don't just present AIC table
- don't take discrete hypotheses too seriously
- be careful

References

Cade, Brian S. 2015. "Model Averaging and Muddled Multimodel Inference." *Ecology*, March. doi:10.1890/14-1639.1.

Fletcher, David, and Peter W. Dillingham. 2011. "Model-Averaged Confidence Intervals for Factorial Experiments." *Computational Statistics & Data Analysis* 55 (11): 3041–48. doi:10.1016/j.csda.2011.05.014.

Harrell, Frank. 2001. *Regression Modeling Strategies*. Springer.

McGill, Brian J. 2015. "Why AIC Appeals to Ecologist's Lowest Instincts." *Dynamic Ecology*. https://dynamicecology.wordpress.com/2015/05/21/why-aic-appeals-to-ecologists-lowest-instincts/.

O'Hara, Bob. 2007. "Focus on DIC." *Deep Thoughts and Silliness*. http://deepthoughtsandsilliness.blogspot.com/2007/12/focus-on-dic.html.

Richards, Shane A. 2005. "Testing Ecological Theory Using the Information-Theoretic Approach: Examples and Cautionary Results." *Ecology* 86 (10): 2805–14. http://www.jstor.org/stable/3450706.

Ripley, Brian D. 2004. "Selecting Amongst Large Classes of Models." *Methods and Models in Statistics: In Honor of Professor John Nelder, FRS*, 155–70. http://books.google.com/books?hl=en&lr=&id=Oulpj4mTi6UC&oi=fnd&pg=PA155&dq=%22expect+to+%EF%AC%81nd,+for+example+linear+models,+AR(p)%22+%22main+distinction+I+would+draw+is+between+explanation+and%22+%22Nowadays+many+experiments+are+done+with+microarrays+to+%EF%AC%81nd%22+&ots=-ZbILPdnqB&sig=yII68OiwIFyK0OrJegoPxhCMc4o.

Schielzeth, Holger. 2010. "Simple Means to Improve the Inter-

pretability of Regression Coefficients." *Methods in Ecology and Evolution* 1: 103–13. doi:10.1111/j.2041-210X.2010.00012.x.

Spiegelhalter, D. J., N. Best, B. P. Carlin, and A. Van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society B* 64: 583–640.

Vaida, Florin, and Suzette Blanchard. 2005. "Conditional Akaike Information for Mixed-Effects Models." *Biometrika* 92 (2): 351–70. doi:10.1093/biomet/92.2.351.

Yang, Yuhong. 2005. "Can the Strengths of AIC and BIC Be Shared? A Conflict Between Model Indentification and Regression Estimation." *Biometrika* 92 (4): 937–50. doi:10.1093/biomet/92.4.937.