Model Selection

Information-Theoretic Approach

UF 2015 (25 minutes)

Outline:

- Why use model selection
- AIC
- AIC weights and model averaging
- Other methods

USGS

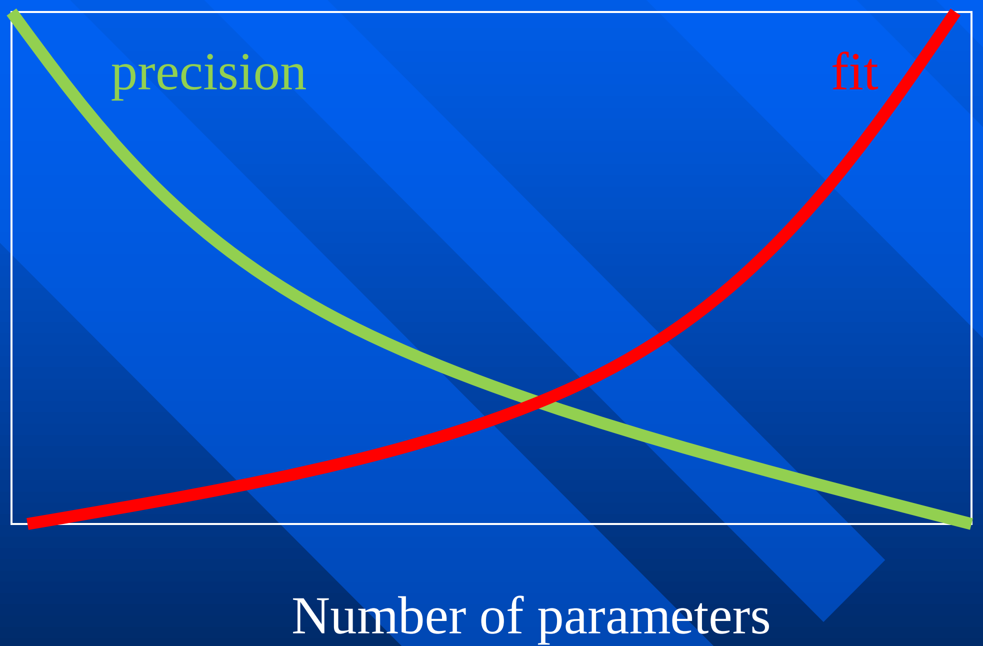
science for a changing world

PATUXENT Wildlife Research Center

Data-Based Model Selection

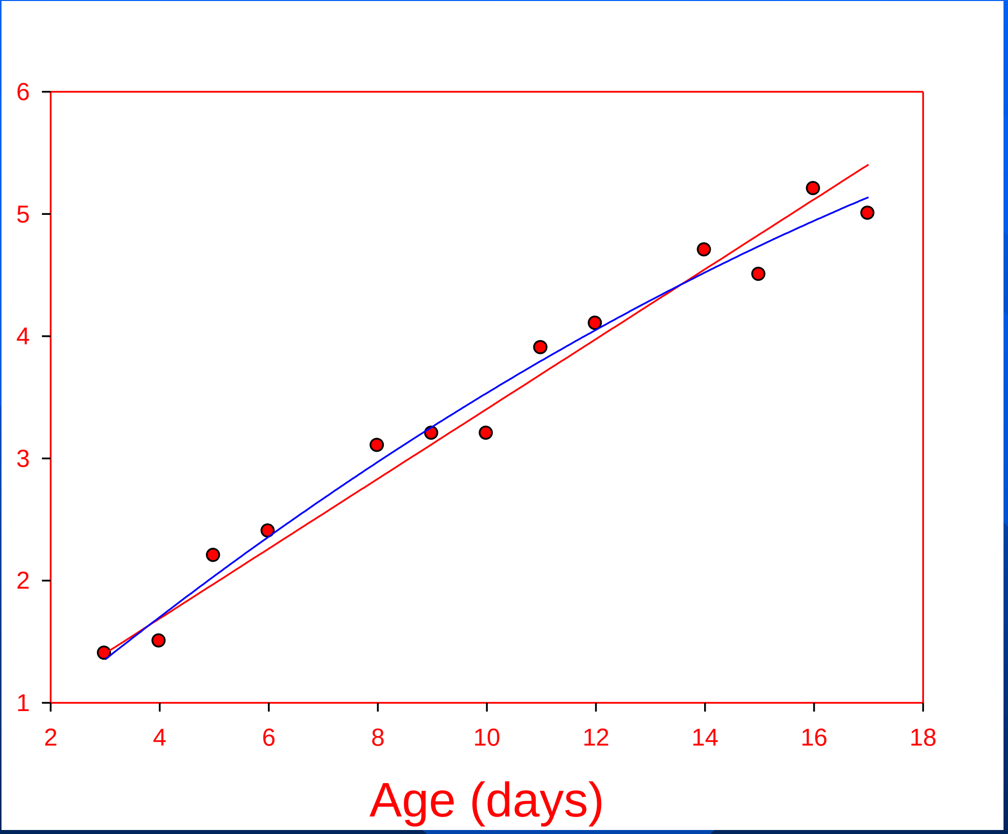
- Problem
  - Multiple plausible models and a single data set
  - How does one select the most reasonable and useful model
- Guiding Principle: "Principle of Parsimony"
  - General trade-off between model fit and estimator precision

Trade off between fit & precision



Example

Weight (g)



$$Y = \beta 0 + \beta 1\, x + \beta 2\, x^2$$

Approaches to Model Selection: Sequential Hypothesis Testing

- For "nested" models
- Begin with most general model and test against the next most general model, etc., down to the simplest model
- Test less general models (H_0) against more general models (H_a) using, e.g., LRT
- If test is "significant", then the extra parameters of H_a are deemed necessary to explain the data
- If test is not "significant", then select the less general model, as it will yield smaller variances (fewer parameters); Principle of Parsimony

USGS
science for a changing world

PATUXENT Wildlife Research Center

Akaike's Information Criterion, AIC

$$AIC = -2\log[L(\hat{\theta} \mid y)] + 2K$$

$L(\hat{\theta} \mid y)$ =likelihood function evaluated at MLEs of θ given the data, *y*

K = number of model parameters

Akaike (1973), Burnham and Anderson (1998, 2002)

USGS
science for a changing world

Quasilikelihood Adjustment for Lack of Fit

- When most general model in model set does not fit data, quasilikelihood procedures are used to adjust tests and model selection metrics for lack of fit caused by overdispersion
- Quasilikelihood variance inflation factor:

$$\hat{c} = \chi^2_{GOF} / df$$

USGS
science for a changing world

PATUXENT Wildlife Research Center

Quasilikelihood Adjusted AIC, QAIC

$$QAIC = -2\log[L(\hat{\theta} \mid y)]/\hat{c} + 2K$$

- Favor simpler models

science for a changing world

AIC Adjusted for Overdispersion *and* Small Sample Size

$$QAIC = -2\log[L(\hat{\theta} \mid y)]/\hat{c} + 2K + \frac{2K(K+1)}{n-K-1}$$

K = number of parameters
n = sample size (e.g., number of releases in CR modeling)

AIC Weights

- w_i = AIC weights ~ weight of evidence in favor of model *i* being most appropriate, given the data and the model set (*R* models)
- Δ_i = $\text{AIC}_i - \text{AIC}_{\min}$ = difference between AIC for model *i* and lowest AIC

$$w_i = \frac{\exp(-\Delta_i / 2)}{\sum_{m=1}^{R} \exp(-\Delta_m / 2)}$$

USGS science for a changing world

PATUXENT Wildlife Research Center

Model Averaging: Incorporating Model Uncertainty

$$\hat{\theta} = \sum_{i=1}^{R} w_i \hat{\theta}_i$$

$$\hat{\text{var}}(\hat{\theta}) = \left[\sum_{i=1}^{R} w_i \sqrt{\hat{\text{var}}(\hat{\theta}_i \mid M_i) + (\hat{\theta}_i - \hat{\theta})^2} \right]$$

$\hat{\theta}_i$ = parameter estimate from model *i*

$\hat{\text{var}}(\hat{\theta}_i \mid M_i)$ = model-specific sampling var

Cooch and White (2015)

model	*estimate*	*variance*	*AIC weight*
$\varphi_c p_t$	0.90888	0.007321	0.85650
$\varphi_c p.$	0.70714	0.002447	0.13345
$\varphi_t p.$	0.74342	0.002343	0.00859
$\varphi_t p_t$	0.88112	0.012090	0.00124
$\varphi_{c*t} p.$	0.74464	0.002332	0.00020
$\varphi_{c*t} p_t$	0.88220	0.011903	0.00002

The model averaged value for the parameter is $\hat{\bar{\varphi}} = 0.88047$. So,

$$
\begin{aligned}
\widehat{\text{var}}(\hat{\bar{p}}_{2,p}) &= \sum_{i=1}^{R} w_i \left[\widehat{\text{var}}(\hat{\theta}_i \mid M_i) + (\hat{\theta}_i - \hat{\bar{\theta}})^2 \right] \\
&= 0.85650\left[0.007321 + (0.90888 - 0.88047)^2\right] \\
&\quad + 0.13345\left[0.002447 + (0.70714 - 0.88047)^2\right] \\
&\quad + \cdots \\
&\quad + 0.00002\left[.011903 + (0.88220 - 0.88047)^2\right] \\
&= 0.011498
\end{aligned}
$$

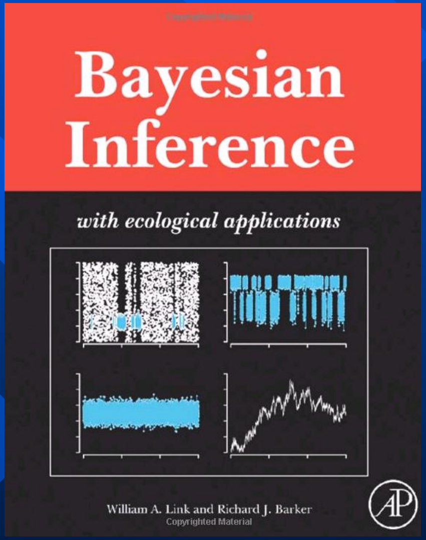
So, our estimate of the *unconditional* variance of the encounter probability p_2 for the poor colony is 0.011498. The standard error is estimated simply as the square-root of the variance: $\sqrt{0.011498} = 0.10723$,

USGS science for a changing world

PATUXENT Wildlife Research Center

Other model selection criterion

- Other model selection criterion: BIC
- See Cooch and White (2015) ; Link and Barker (2010)
- Active area of research



Take home points

- trade-off between model fit and estimator precision
- $AIC = -2\log[L(\hat{\theta} \mid y)] + 2K$
- Adjustment for sample size & overdispersion
- AIC weight
- Model Averaging