Model-Averaged Confidence Intervals

David Fletcher
Department of Mathematics and Statistics
University of Otago
Dunedin
New Zealand

June 18, 2013

PhD Thesis Work: Daniel Turek and Jimmy Zeng

Overview

- Model Uncertainty and Model Averaging
- Model-averaged confidence intervals
  - Wald
  - Model-averaged tail area (MATA)
  - Simulation study
- Related work
  - Profile likelihood version of MATA interval
  - Bootstrap version of MATA interval
  - Kullback-Leibler prior for Bayesian model-averaging

Model Uncertainty and Model Averaging

- Single-model inference
- Model selection
- Use of "best model"
  - Bias
  - Underestimation of uncertainty
- Bayesian Model Averaging
  - Natural
  - Reversible Jump MCMC
- Frequentist Model Averaging
  - Weighted mean of estimates from each model
  - Weight based on an information criterion (or bootstrap)

Notation

Reponse variable Y and predictor variables X

Prediction: $\theta = E(Y|X = x)$

Model set: $\{M_1, M_2, \ldots, M_R\}$

Model-averaged estimate

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

w_i = weight for M_i (based on AIC, AICc or BIC)

$\hat{\theta}_i$ = estimate under M_i

Analogous to mean of Bayesian model-averaged posterior for θ

Model-Averaged Confidence Interval

Wald interval (Burnham and Anderson 2002)

$$\bar{\theta} \pm z \left[\sum_{i=1}^{R} w_i \left\{ \left(\frac{t_{\nu_i}}{z} \right)^2 \hat{V}\left(\hat{\theta}_i\right) + \left(\hat{\theta}_i - \bar{\theta}\right)^2 \right\} \right]^{1/2}$$

$z =$ percentile of $N(0, 1)$

$t_{\nu_i} =$ percentile of t-distribution with ν_i d.f.

$\nu_i =$ error d.f. for M_i

Model-Averaged Confidence Interval

Wald interval (Burnham and Anderson 2002)

$$\bar{\theta} \pm z \left[\sum_{i=1}^{R} w_i \left\{ \left(\frac{t_{\nu_i}}{z} \right)^2 \hat{V}\left(\hat{\theta}_i\right) + \left(\hat{\theta}_i - \bar{\theta}\right)^2 \right\} \right]^{1/2}$$

Non-normal distribution of $\bar{\theta}$ (even if we transform θ)

Estimation of standard error (weights are random variables)

Alternative: Claeskens and Hjort 2008

Model-Averaged Confidence Interval

Model-averaged tail area (MATA) interval

Upper limit is the solution of

$$\sum_i w_i F_i(t_i) = \alpha$$

$t_i = \frac{\hat{\theta}_i - \theta}{\sqrt{\hat{V}(\hat{\theta}_i)}}$

$Y \sim$ Normal: $F_i(\cdot) =$ c.d.f. of t-distribution with ν_i d.f.

$\hat{\theta}_i \sim$ Normal: $F_i(\cdot) =$ c.d.f. of $N(0, 1)$

Lower limit uses $1 - F_i(t_i)$ in place of $F_i(t_i)$

Model-Averaged Confidence Interval

Model-averaged tail area (MATA) interval

Upper limit is the solution of

$$\sum_i w_i F_i(t_i) = \alpha$$

- Acceptance region using weighted p-value
- Analogous to Bayesian model-averaged credible interval

Simulations

Normal linear regression: $y_i \sim N(\mu_i, \sigma^2)$

$$M_1\colon\ \mu_i = \beta_0$$
$$M_2\colon\ \mu_i = \beta_0 + \beta_1 x_{1i}$$
$$M_3\colon\ \mu_i = \beta_0 + \beta_2 x_{2i}$$
$$M_4\colon\ \mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}$$
$$M_5\colon\ \mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i}$$

$\sigma = 1,\ \beta_0 = 1,\ \beta_1 = 0.3,\ \beta_2 = 0.3,\ \beta_{12} = 0.1$

$x_{1i} \sim \text{Normal}(0, 1),\ x_{2i} \sim \text{Gamma}(2, 1)$

$n = 15, 20, \ldots, 100$

Simulations

Prediction of μ

For 25 combinations of percentiles of x_1, x_2

Show results for 50^{th} percentile of x_1, 90^{th} percentile of x_2

Generating model

Largest (M_5)

Randomly selected: $Pr(M_i) = 1/5 \quad (i = 1, \ldots, 5)$

Model weights: $w_i \propto exp(-0.5IC_i), \quad IC_i = -2\log\hat{L}_i + kp_i$

AIC $\Rightarrow k = 2$

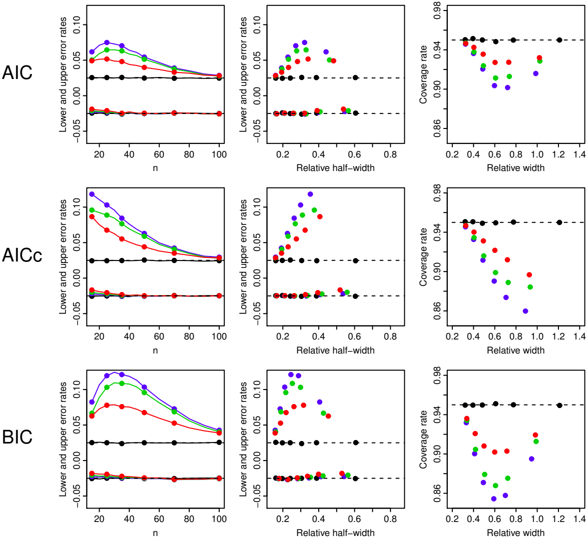
AICc $\Rightarrow k = 2n/(n - p - 1)$

BIC $\Rightarrow k = \log n$

Relative half-widths: $\frac{\theta - \theta_L}{\theta}$ and $\frac{\theta_U - \theta}{\theta}$

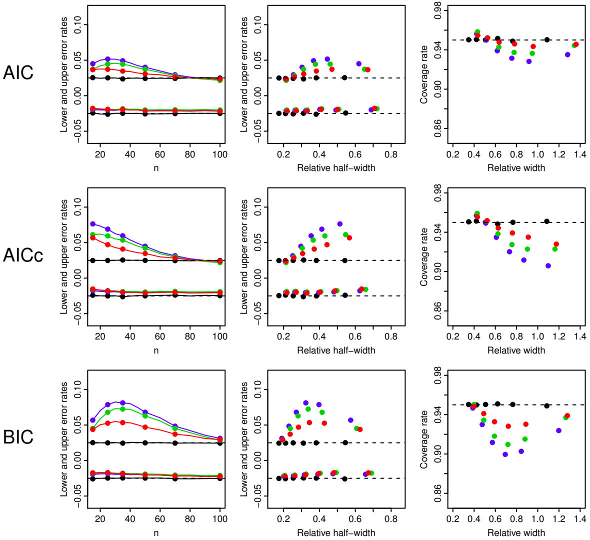
Generating model: Largest

Largest, MATA, Wald, Wald (alternative SE formula)



Generating model: Randomly selected

Largest, MATA, Wald, Wald (alternative SE formula)



Conclusions and Discussion

- Largest differences for upper error rate
- MATA better than Wald
- AIC weights best (c.f. Fletcher and Dillingham 2011)

- Profile likelihood version (Fletcher and Turek 2011)
- Bootstrap version (Zeng, Fletcher and Dillingham, submitted)
- Kullback-Leibler prior (Turek, submitted)