Bayesian model selection

Modelos Bayesianos con aplicaciones ecológicas
Dr. Cole Monnahan
University of Concepción, Chile
Enero, 2018

Recap

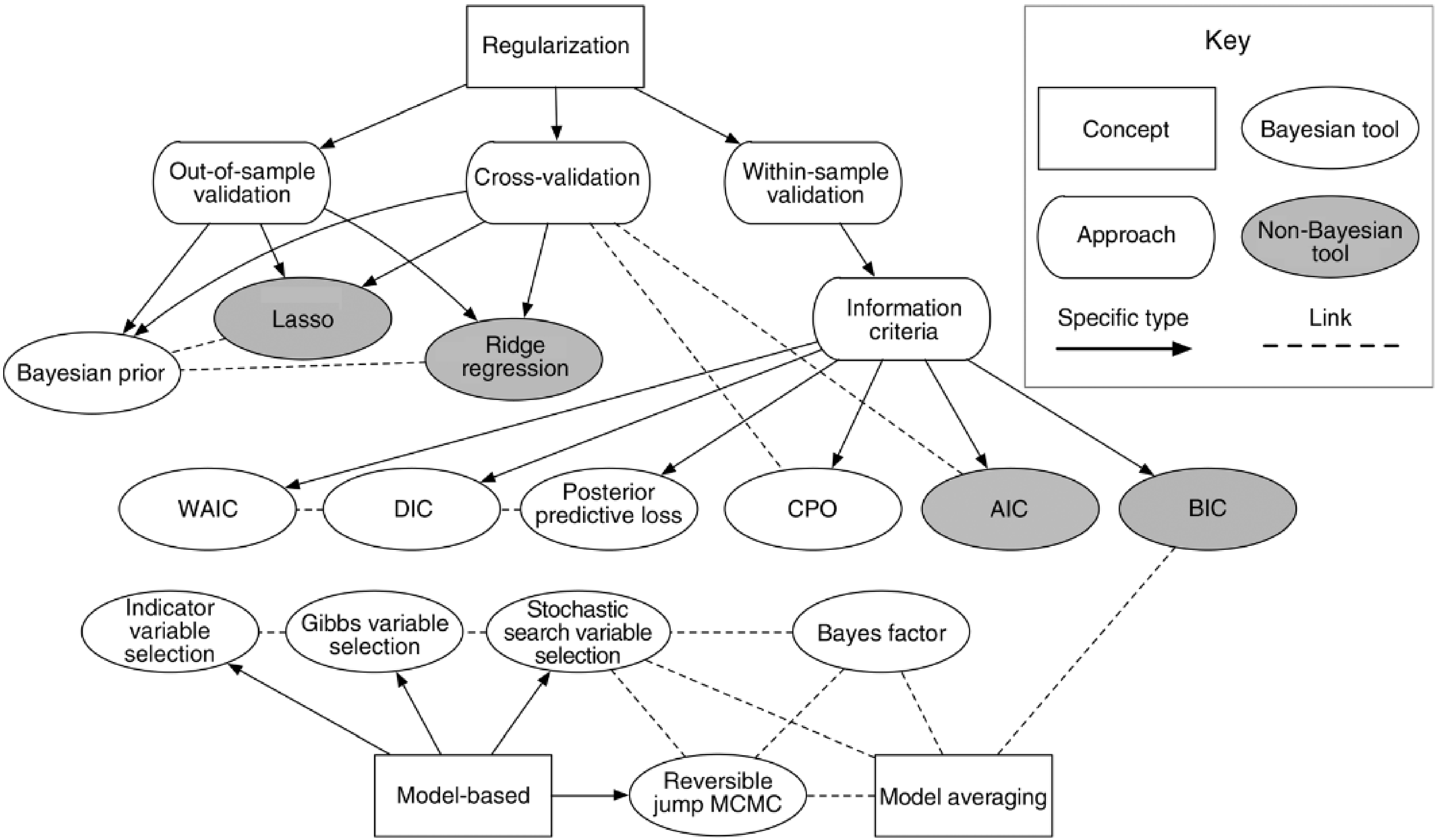
- Usamos *posterior predictive distributions* para probar “model self-consistency”
- El modelo tuvo un dato que no era bueno
- Quizas con extra estructura con covariables el modelo puede explicar los datos
- Vamos a usar *posterior predictive distribution* y DIC

Bayesian model selection

- Model selection is not trivial. There are many existing tools and more being developed.
- According to Hobbs & Hooten (2018)
  1. Out of sample is the best approach if possible
  2. K-fold cross validation good but slow
  3. DIC is good when prediction is important and model is slow, works best when (# pars << # data)
  4. WAIC good for hierarchical models

Hobbs & Hooten (2018)

Bayesian model selection



Hobbs & Hooten (2018)

Tarea

- We will try 3 versions of our binomial survival model (GLM)
- Fit3= de lectura (problema con sitio 5)
- Fit4= con solo covariable x1
- Fit5= con covariables x1 y x2

```
x1 <- c(9.450, 8.079, 7.686, 8.003, 2.882, 11.095, 10.696, 8.263, 12.043,
        9.238)
x2 <- c(0.08, 0.252, 0.158, 0.25, 0.081, 0.037, 0.002, 0.042, 0.053,
        -0.02, 0.141, -0.001, -0.078, 0.103, -0.076, 0.037, 0.071, 0.113,
        0.23, 0.246)
```

Assume p= ilogit(theta + beta1*x1 + beta2*x2)

Tarea

- Fit the 3 versions of the model (use separate .jags files)
- Recreate posterior predictive plots for all three
- Which model does DIC select? (`fit3$BUGSoutput$DIC` etc.)
- [What should you use for priors on coefficients?]
- [Suggestion: Normalize the covariates]

Tarea

- Use the results to select which model you think is most appropriate.
- Write down the model, likelihood and priors.
- Calculate the probability of survival greater than 0.8. P(p>0.8) from this model