Bayesian model selection

Modelos Bayesianos con aplicaciones ecológicas

Dr. Cole Monnahan
University of Concepción, Chile

Enero, 2019

Recap

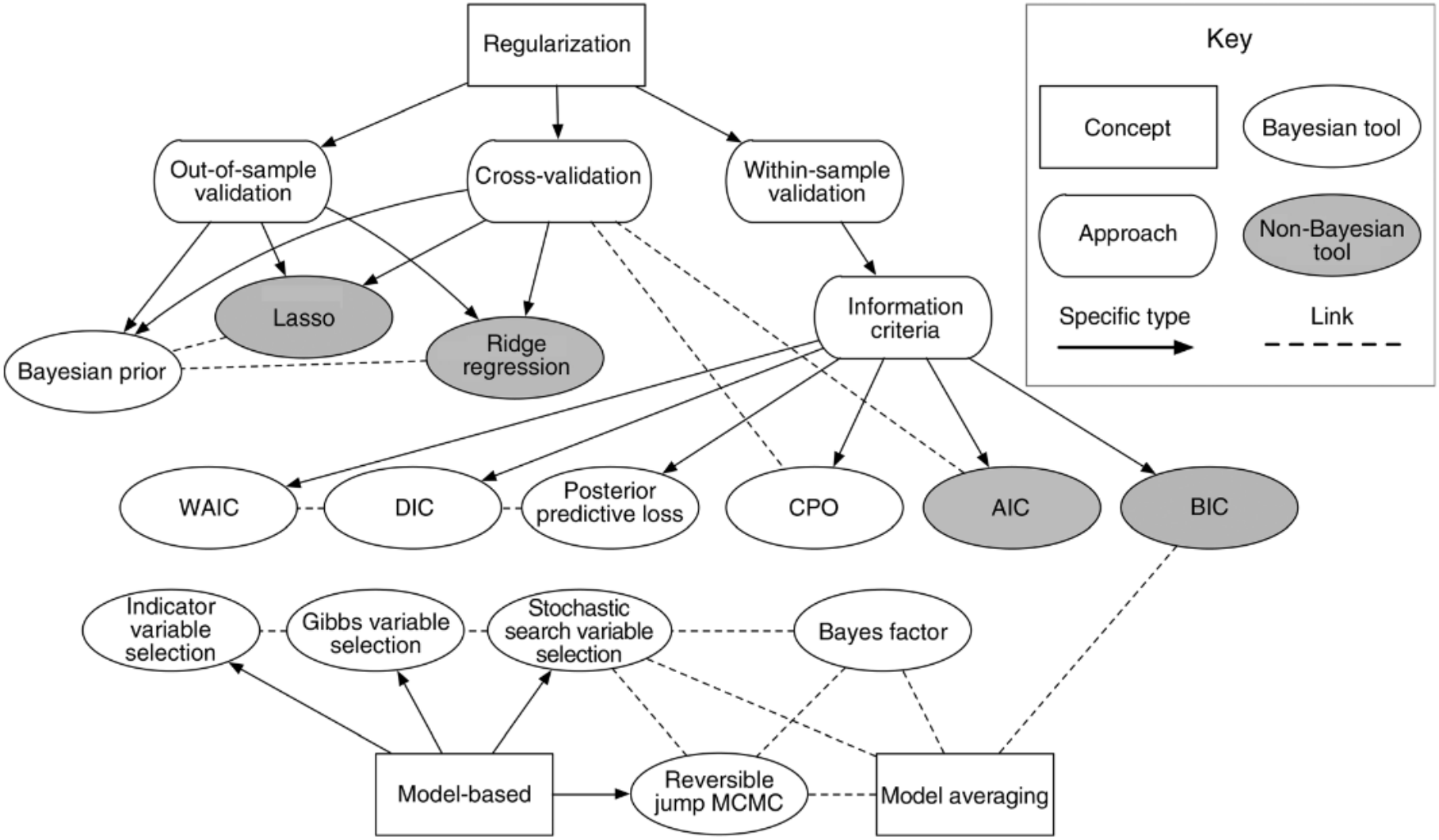
- Usamos *posterior predictive distributions* para probar “model self-consistency”
- El modelo tuvo un dato que no era bueno
- Quizás con extra estructura con covariables el modelo puede explicar los datos
- Vamos a usar *posterior predictive distribution* y DIC

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 2 versions of our binomial survival model (GLM)
- Fit3= de lectura (problema con sitio 5)
- Fit4= con covariable x1

```
x1 <- c(9.450, 8.079, 7.686, 8.003, 2.882, 11.095, 10.696, 8.263, 12.043,
        9.238)
```

Assume p= ilogit(theta + beta1*x1)

Tarea

- Fit the 2 versions of the model (create logistic 4.jags)
- Recreate posterior predictive plots for both
- 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.
- Calculate the probability of survival greater than 0.8, P(p>0.8), for site 3

- Read the paper http://www.scielo.org.mx/scielo.php?pid=S0185-38802009000200005&script=sci_arttext&tlng=pt
- And write down:
  - What is the type of model?
  - What is the likelihood function?
  - What are the priors?