Bayesian inference

Assessment workshop

Arni Magnusson

ICES, 18–22 Feb 2013

Outline

1. Overview - background, inference, priors, MCMC

2. Application - practical Bayesian, MCMC, exercise

Outline

1. Overview - background, inference, priors, MCMC
2. Application - practical Bayesian, MCMC, exercise

History

Bayes (1763) and Laplace (1774)

Bayesian inference with uniform priors

Fiercely debated in the 1930s–1960s

Dismissed by Fisher and Neyman

Widely used since the 1990s

Gelman et al. (1995), MCMC

http://cran.r-project.org/web/views/Bayesian.html

Bayesian theory

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

$$P(\theta|Y) = \frac{P(Y|\theta) \times P(\theta)}{P(Y)}$$

$$P(\theta|Y) \propto L(\theta|Y) \times P(\theta)$$

$$f = -\log L + \text{Penalty}$$

Statistical inference

Significance tests, p values

Given that the null hypothesis is true . . . probability of getting a sample. . .
Sometimes all we need

Uncertainty, confidence intervals

Given that we know the true parameter values . . .
. . . this is the distribution of estimates in repeated experiments . . .

Subjective probability

Probability redefined to evaluate $P(\theta|Y)$ instead of $P(Y|\theta)$

Likelihood inference, uniform prior

Support, penalized likelihood
$f = -\log L + \text{constant}$

Prior distribution

- Uniform
- Conjugate
- Objective, empirical, meta analysis
- Assist model convergence without fixing a parameter

Markov chain Monte Carlo

Metropolis algorithm

Start from the best fit, and then:

1. Sample proposal $\theta^\star$ from jumping distribution at time t, $J_t(\theta^\star \mid \theta^{t-1})$
2. Calculate the ratio of the densities, $r = \frac{p(\theta^\star \mid y)}{p(\theta^{t-1} \mid y)}$
3. Set $\theta^t = \begin{cases} \theta^\star & \text{with probability } \min(r, 1) \\ \theta^{t-1} & \text{otherwise} \end{cases}$

Repeat *many* times

Markov chain Monte Carlo

Metropolis-Hastings algorithm

Extends the Metropolis algorithm:

- Jumping distribution $J_t(\theta^\star \mid \theta^{t-1})$ can be asymmetric
- To accommodate the asymmetry, the ratio becomes

$$r = \frac{p(\theta^\star|y) \,/\, J_t(\theta^\star|\theta^{t-1})}{p(\theta^{t-1}|y) \,/\, J_t(\theta^{t-1}|\theta^\star)}$$

Outline

1. Overview - background, inference, priors, MCMC
2. Application - practical Bayesian, MCMC, exercise

In practice

Frequentist and Bayesian methods address different questions

How likely is $\theta = 160$ or $\theta = 220$?
What is the probability of $B < B_{\text{lim}}$?

MCMC is practical

Evaluate uncertainty
Find global minimum
Diagnose model behavior

Exercise

Use a prior for the Beverton-Holt parameter $R_{\max}$

$$R \;=\; R_{\max} \frac{S}{S + S_{50}}$$

Imagine that we are having model convergence problems, but an ecosystem model has indicated that asymptotic recruitment for this stock might be around 1600.

Use normal prior distribution with $\mu = 1600$. Try (1) $\sigma = 500$, (2) $\sigma = 100$, and (3) shrinking the dataset down to the first $n = 30$ observations.

Hint: $f \;=\; -\log L \;+\; \text{Penalty}$ in (3), set `n` in the `.dat` file