Chapter 10

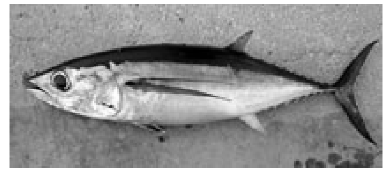
Bayesian Diagnostics

- Convergence diagnostics.
- Posterior predictive checks.
- DIC, model selection, and complexity.
- Bayes factors
- Sensitivity analysis

Convergence diagnostics

- Primarily, to assess whether the MCMC chain has converged to a stationary distribution.
- We will use the CODA package in R. This implements
  - Gelman-Rubin diagnostic, based on multiple chains.
  - Geweke diagnostic for stationarity.
  - Heidelberger-Welch stationarity and run-length diagnostics.
  - Raftery-Lewis run-length diagnostic (for quantiles).
- The CODA package also contains:
  - Diagnostic plots
  - Functions for manipulating BUGS output
  - A function to find HPD intervals

State-space Schaefer surplus prodn model
(Millar and Meyer, 2000)



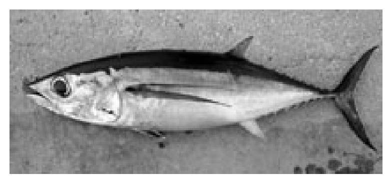
Observation equation: $y_t = qB_t e^{v_t}$ where v_t are iid Normal(0,τ^2).

Process equations:
$$\begin{aligned} B_1 &= K e^{u_1} &&, \quad t = 1 \\ B_t &= \left(B_{t-1} + rB_{t-1}(1 - B_{t-1}/K) - C_{t-1}\right) e^{u_t} &&, \quad t > 1 , \end{aligned}$$

where u_t are iid Normal(0,σ^2).

Priors:
$$\log(q) \sim U(-\infty,\infty), \quad \text{i.e., } \pi(q) = 1/q$$
$$\log(K) \sim N(5.04, 0.516^2)$$
$$\log(r) \sim N(-1.38, 0.51^2)$$
$$\tau^{-2} \sim \Gamma(1.71, 0.00861)$$
$$\sigma^{-2} \sim \Gamma(3.79, 0.0102)$$

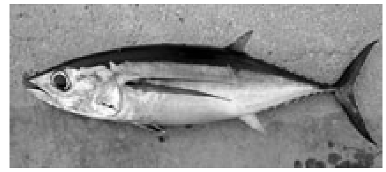
Surplus prodn model ctd: Convergence diagnostics



For this case study, the general steps of the diagnostic checking are:

- Generate 2 chains of length 250,000 and thinned by a factor of 25 (i.e., save every 25th value).
- Save index and output to text files.
- Read into R
- Perform convergence (i.e., stationarity) tests.

Surplus prodn model ctd: Convergence diagnostics



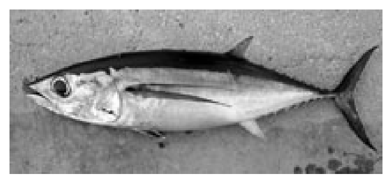
```
library(coda)
Chain1=read.coda("Pars1.out","Pars.ind")
Chain2=read.coda("Pars2.out","Pars.ind")

#Convergence diagnostics
geweke.diag(Chain1); geweke.diag(Chain2)
heidel.diag(Chain1); heidel.diag(Chain2)
raftery.diag(Chain1); raftery.diag(Chain2)
gelman.diag(mcmc.list(Chain1,Chain2))

#Find 95% intervals of highest posterior density
#Separately
HPDinterval(mcmc.list(Chain1,Chain2),prob=0.95)
#Both chains combined
HPDinterval(mcmc(rbind(Chain1,Chain2)),prob=0.95)

#The above can also be done using the interactive codamenu() command
```


Surplus prodn model ctd: Posterior predictive checks



Millar and Meyer (2000) performed four different posterior predictive tests, within R.

See R code provided.

DIC, model selection, and complexity.

Spiegelhalter et al. (2002) formalized the concept of the deviance information criterion, DIC, as a measure of model fit and complexity.

The deviance is -2 times the log-likelihood

$$D(\theta) = -2\log(f(y \mid \theta))$$

Define "Dbar" as the posterior mean of the deviance

$$\overline{D(\theta)} = E_{\theta|y}\left[D(\theta)\right]$$

and "Dhat" as the deviance evaluated at some plug-in estimate of θ, typically the posterior mean of θ.

$$\hat{D} = D(\bar{\theta}) = D\left(E_{\theta|y}[\theta]\right)$$

DIC, model selection, and complexity.

Using somewhat heuristic arguments, Spiegelhalter et al. (2002) argued that

$$p_D = \overline{D(\theta)} - D(\bar{\theta})$$

quantifies the "effective number of parameters" in the model.

Model goodness of fit is a trade-off between model fit and model complexity. The DIC is defined to be

$$DIC = D(\bar{\theta}) + 2p_D$$

or equivalently,

$$DIC = \overline{D(\theta)} + p_D$$

DIC, model selection, and complexity.

Spiegelhalter et al. (2002) developed DIC in the context of hierarchical models, but it has since been applied much more widely.

However, there is concern amongst some in the Bayesian community that the DIC is not reliable and is being over-used.

Another measure of model complexity is given by half the variance of the deviance

$$p_D^{(2)} = 0.5\,\mathrm{var}_{\theta|y}\left(D(\theta)\right)$$

Stock-recruit example

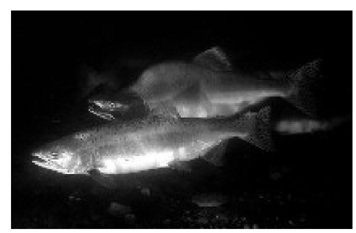
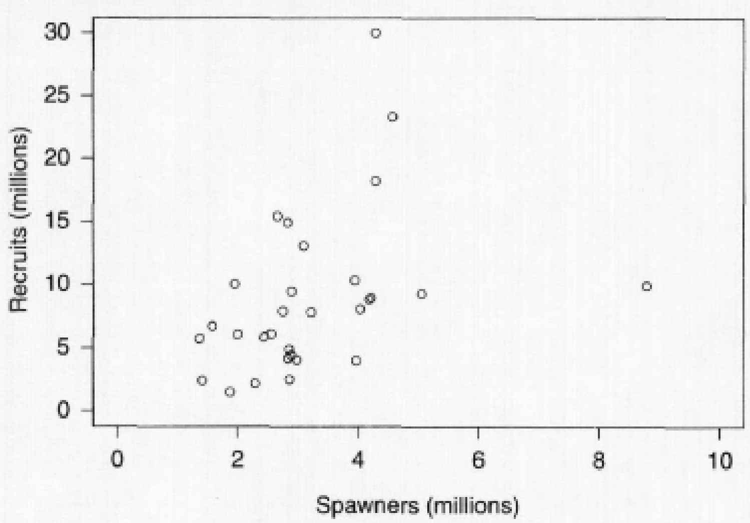
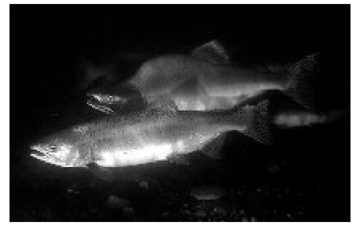


Fig. 1. Alaska pink salmon (*Oncorhynchus gorbuscha*) escapement (spawners) and recruitment data from p. 105 of Quinn and Deriso (1999).



Stock-recruit example



Use DIC to compare fits of the Ricker, and Beverton-Holt, stock-recruit curves to these data.

Bayes factors

To choose between two models, the Bayes factor is the ratio of the marginal densities for the data under the two models:

$$B_{12} = f_1(y) / f_2(y)$$

where

$$f_i(y) = \int_{\theta_i} f_i(y \mid \theta_i) \pi_i(\theta_i) d\theta_i$$

Bayes factors: calculation

Bayes factors are challenging to calculate (e.g., see Kass and Raftery (1995)) and hence are not routinely used.

Recently, Brodziak and Legault (2005) used Bayes factors to average over stock-recruitment curves, in the context of estimating rebuilding targets. They calculated approximate Bayes factors based on the Schwarz information criterion.

The Schwarz criterion requires maximization of the model likelihood, hence Brodziak and Legault (2005) used the Automatic Differentiation Model Builder software.

Sensitivity analysis

A seminar on this topic is scheduled for the afternoon of Tuesday 9 May.