Chapter 8 Hierarchical Models

In many situations, there will be data from several/many subpopulations or groups.

Rather than performing separate analyses for each group, it may make good sense to assume that there is some relationship between the parameters of different groups. That is, introduce a higher level of randomness on the parameters.

This is a meta-analysis approach, and is a particularly effective approach when the information from each sub-population is limited.

Hierarchical Models

Hierachical modeling also includes mixed-effects models, variance component models and continuous mixture models.

The frequentist application of meta-analysis is often mis-called empirical Bayes modeling.

Meta-analysis Example 1

Myers et al. (1999, Can. J. Fish. Aquat. Sci., 56: 2404-2419) perform frequentist meta-analyses of maximum population reproductive rate, for a wide variety of species.

They fitted the Ricker stock-recruit curve

$$R_t = \alpha S_t e^{-\beta S_t}$$

In the case of a semelparous species, α is the maximum population reproductive rate.

Meta-analysis Example 1

Lognormal error on R_t was assumed. That is

$$y_t = \log(R_t / S_t) = \log(\alpha) - \beta S_t + \varepsilon_t$$

where ε_t are iid N(0,σ^2).

Letting i index the sub-populations, our notation is

$$y_{i,t} = \log(R_{i,t} / S_{i,t}) = \log(\alpha_i) - \beta_i S_{i,t} + \varepsilon_{i,t}$$

where $\varepsilon_{i,t}$ are iid N(0,σ_ι^2).

Meta-analysis Example 1, ctd…

Myers et al. assumed that α_i's were log-normally distributed across the different sub-populations of each species.

That is, for sub-population i

$$\log \alpha_i \sim N(\mu, \tau^2)$$

No relationship between β_i's was assumed.

The hierarchical model requires hyper-priors on parameters μ and τ^2.

What would be a good choice for the hyper-priors?

Meta-analysis Example 1, ctd…

$$\log \alpha_i \sim N(\mu, \tau^2)$$

The reference hyper-priors for hierarchical models (Gelman, 2005) are

$$\mu \sim Unif(-\infty, \infty)$$

$$\tau \sim Unif(0, \infty)$$

That is, the prior $\pi(\mu, \tau)$ is flat.

Previously, the prior $\pi(\tau) = 1/\tau$ had been used extensively in hierarchical models. However, Gelman (2005) showed this to have undesirable properties.

The WinBUGS 1.4.1 patch modified the hierarchical model examples to use Gelman's hyper-priors.

Meta-analysis Example 1, ctd…

We will fit a Bayesian version of this model to Chinook salmon data, using WinBUGS.

Note that the model also requires priors for each of the β_i and σ_i parameters.

The full prior specification used is:

$$\mu \sim Unif(-\infty, \infty)$$

$$\tau \sim Unif(0, \infty)$$

$$\beta_i \sim Unif(-\infty, \infty)$$

$$\log(\sigma_i) \sim Unif(-\infty, \infty)$$

Meta-analysis Example 2

Millar and Methot (2002) implemented a hierarchical model for the bulk catchability, Q, of U.S. West Coast rockfish (*Sebastes* spp.)
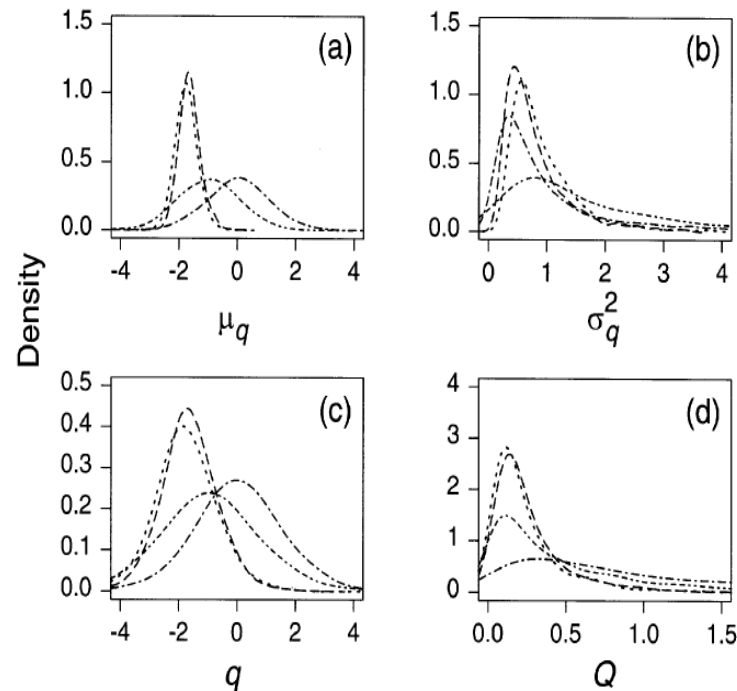
$$q_i = \log(Q_i) \sim N(\mu_q, \sigma_q^2)$$

where vaguely informative hyper-priors were placed on μ_q and σ_q^2.

The model was fitted using Automatic Differentiation Model Builder (ADMB).

Meta-analysis Example 2

Fig. 2. Marginal prior and posterior densities for (a) μ_q, (b) σ_q^2, (c) log(catchability), q, and (d) bulk catchability (Q): vague prior (– · – ·); resulting posterior (– – – –); informative prior (– – · ·); resulting posterior (····).

Meta-analysis Example 3

In the previous two examples, the meta-analysis was applied to "raw" data from each of the sub-populations:

- ‾ Chinook stock-recruit data in Example 1
- ‾ Rockfish CPUE in Example 2

In this third example (Gelman et al. 2003), the sub-population data have already been (individually) analyzed, resulting in a table of point estimates and associated std errors.

Meta-analysis Example 3

```
    school estimate    sd
1      A    28.39   14.9
2      B     7.94   10.2
3      C    -2.75   16.3
4      D     6.82   11.0
5      E    -0.64    9.4
6      F     0.63   11.4
7      G    18.01   10.4
8      H    12.16   17.6
```

The estimate, y_i, is a measure of the effectiveness of special coaching programs on SAT test scores.

The question of interest is whether the coaching programs increase SAT score.

Meta-analysis Example 3

To keep things simple, it will be assumed that the estimated std errors are the true std deviations of the estimates. (We could try relaxing this assumption later.)

So, we have

$$y_i \sim N(\mu_i, \sigma_i^2)$$

where σ_i are assumed known.

The hierarchical bit is

$$\mu_i \sim N(\nu, \tau^2)$$

with flat priors on ν and τ.

Exchangeability

Exchangeability is the assumption that (a subset of) the subpopulation parameters can be assumed to be a random sample from a higher-level distribution.

Care must be taken to ensure that exchangeability is a reasonable assumption. E.g.,

- In the stock-recruit meta-analysis, it may be reasonable to assume that the α_i (the maximum population reproductive rates) are exchangeable for a common species.

- However, would it be reasonable to assume that the β_i were exchangeable?