Introduction to hierarchical (mixed-effects) models part II

FW 891

Christopher Cahill

25 September 2023

Quantitative Fisheries Center
MICHIGAN STATE UNIVERSITY

Purpose

A hierarchical model

$$
\begin{aligned}
y_i &\overset{ind}{\sim} p\left(y \mid \theta_i\right) \\
\theta_i &\overset{iid}{\sim} p(\theta \mid \phi) \\
\phi &\sim p(\phi)
\end{aligned}
$$

A hierarchical model

$$\begin{aligned} y_i &\overset{ind}{\sim} p\left(y \mid \theta_i\right) \\ \theta_i &\overset{iid}{\sim} p(\theta \mid \phi) \\ \phi &\sim p(\phi) \end{aligned}$$

- y_i is observed data

A hierarchical model

$$
\begin{aligned}
y_i &\stackrel{ind}{\sim} p\left(y \mid \theta_i\right) \\
\theta_i &\stackrel{iid}{\sim} p(\theta \mid \phi) \\
\phi &\sim p(\phi)
\end{aligned}
$$

- y_i is observed data
- $\theta = (\theta_1, \ldots, \theta_n)$ and ϕ are parameters

A hierarchical model

$$
\begin{aligned}
y_i &\overset{ind}{\sim} p\left(y \mid \theta_i\right) \\
\theta_i &\overset{iid}{\sim} p(\theta \mid \phi) \\
\phi &\sim p(\phi)
\end{aligned}
$$

- y_i is observed data
- $\theta = (\theta_1, \ldots, \theta_n)$ and ϕ are parameters
- Only ϕ has a prior that is set

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y)$$

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y)$$

- Apply Bayes rule

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi) p(\theta, \phi)$$

- Apply Bayes rule

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi)p(\theta, \phi) = p(y \mid \theta)p(\theta \mid \phi)p(\phi)$$

- Apply Bayes rule
- We can also break this joint distribution down into a conditional distribution of θ given ϕ

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi)p(\theta, \phi) = p(y \mid \theta)p(\theta \mid \phi)p(\phi)$$

- Apply Bayes rule
- We can also break this joint distribution down into a conditional distribution of θ given ϕ
    - Do this using conditional probability rules

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi)p(\theta, \phi) = p(y \mid \theta)p(\theta \mid \phi)p(\phi)$$

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi)p(\theta, \phi) = p(y \mid \theta)p(\theta \mid \phi)p(\phi)$$

- We may also care about the marginal posterior of ϕ:

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi) p(\theta, \phi) = p(y \mid \theta) p(\theta \mid \phi) p(\phi)$$

- We may also care about the marginal posterior of ϕ:

$$p(\theta \mid y) = \int p(\theta, \phi \mid y) d\phi$$

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi) p(\theta, \phi) = p(y \mid \theta) p(\theta \mid \phi) p(\phi)$$

- We may also care about the marginal posterior of θ:

$$p(\theta \mid y) = \int p(\theta, \phi \mid y) d\phi$$

- Or instead the marginal distribution of ϕ:

$$p(\phi \mid y) = \int p(\theta, \phi \mid y) d\theta$$

Posterior distribution for hierarchical models

- The joint posterior distribution of interest in hierarchical models is

$$p(\theta, \phi \mid y) \propto p(y \mid \theta, \phi)p(\theta, \phi) = p(y \mid \theta)p(\theta \mid \phi)p(\phi)$$

- We may also care about the marginal posterior of θ:

$$p(\theta \mid y) = \int p(\theta, \phi \mid y) d\phi$$

- Or instead the marginal distribution of ϕ:

$$p(\phi \mid y) = \int p(\theta, \phi \mid y) d\theta$$

- $p(unknowns|knowns) \propto \text{ assumptions you make}$

Summary and outlook

References

- Berliner, L. M. 1996, Hierarchical Bayesian time series models", Maximum Entropy and Bayesian Methods, 15-22.
- Gelman and Hill 2007. Data analysis using regression and multilevel models
- Royle and Dorazio 2008. Hierarchical modeling and inference in ecology.
- Kery and Schaub. 2012. Bayesian Population Analysis using WinBUGS. Chapter 3
- McElreath 2023. Statistical Rethinking. Second Edition, Chapters 2 and 9.
- Punt et al. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES Journal of Marine Sci.

https://www.youtube.com/watch?v=nNQdvXfW73E

Exercises (optional)

1. **Prior sensitivity test and model evaluation**

It is always good to think critically about model fit. Check the sensitivity of your estimates to the priors used in the sripts, and conduct standard model evaluations. Report your findings.

2. **Derived parameters**

A biologist wants to know whether scorpions in populations 2 and 3 are “different” in some way. Kick out the difference between α_{pop_2} and α_{pop_3} as a derived variable. Calculate the $\mathrm{Pr}(\alpha_{pop_2} > \alpha_{pop_3})$ and report this information back to your biologist friend.

3. **Hyperdistributions**

The hyperdistribution contains information on the among-group variability in scorpion weight-length relationship intercept terms. Generate a posterior predictive hyperdistribution for α_{pop} and plot it to examine the uncertainty we might expect if we were to measure scorpions from an entirely new population (but was otherwise similar to our original three populations)