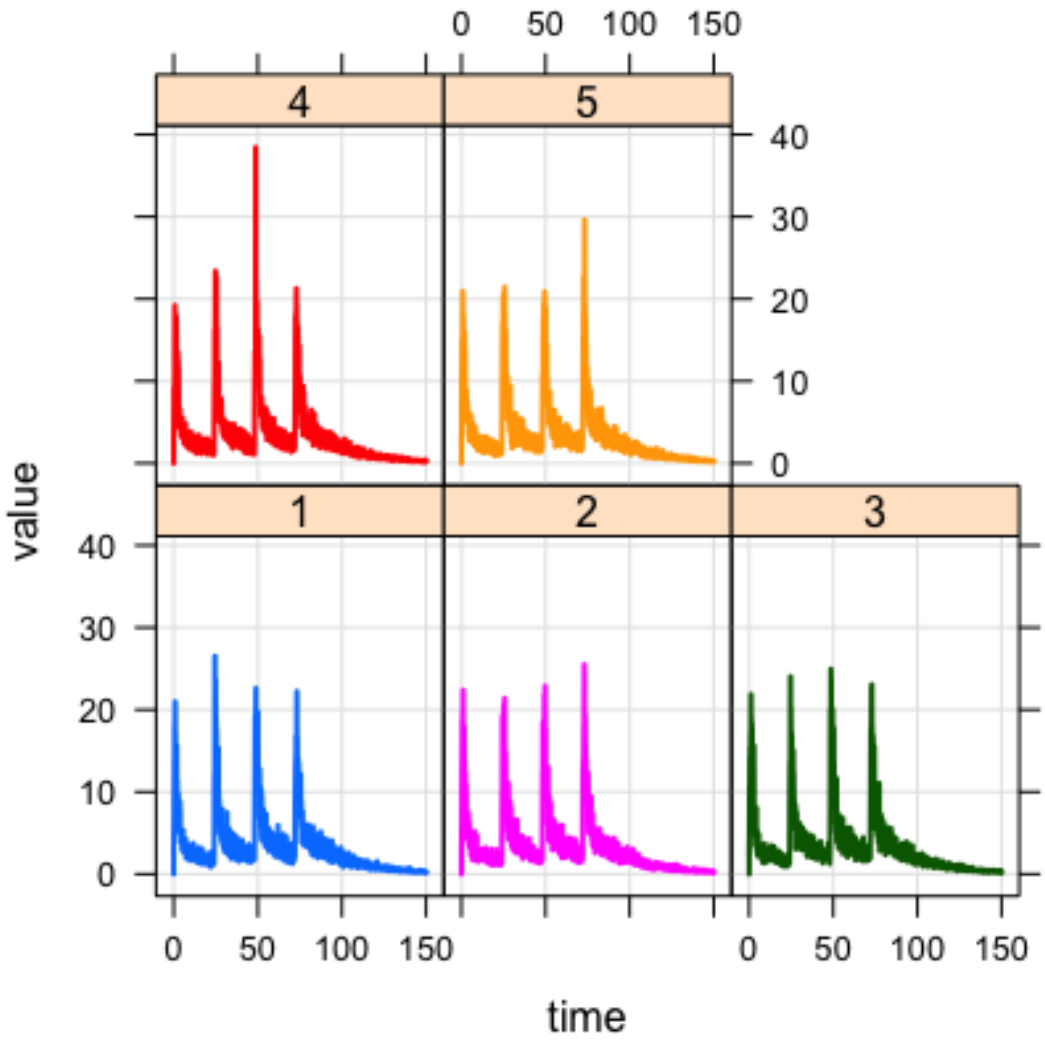
IV

Population models

Suppose our data can be pooled into groups.

- medical measurements are grouped by patients
- sport measurements are grouped by players
- people's voting intention can be grouped by states, social status, etc.



Simulated with *mrgsolve* (https://mrgsolve.github.io/)

With a hierarchical model, we can:

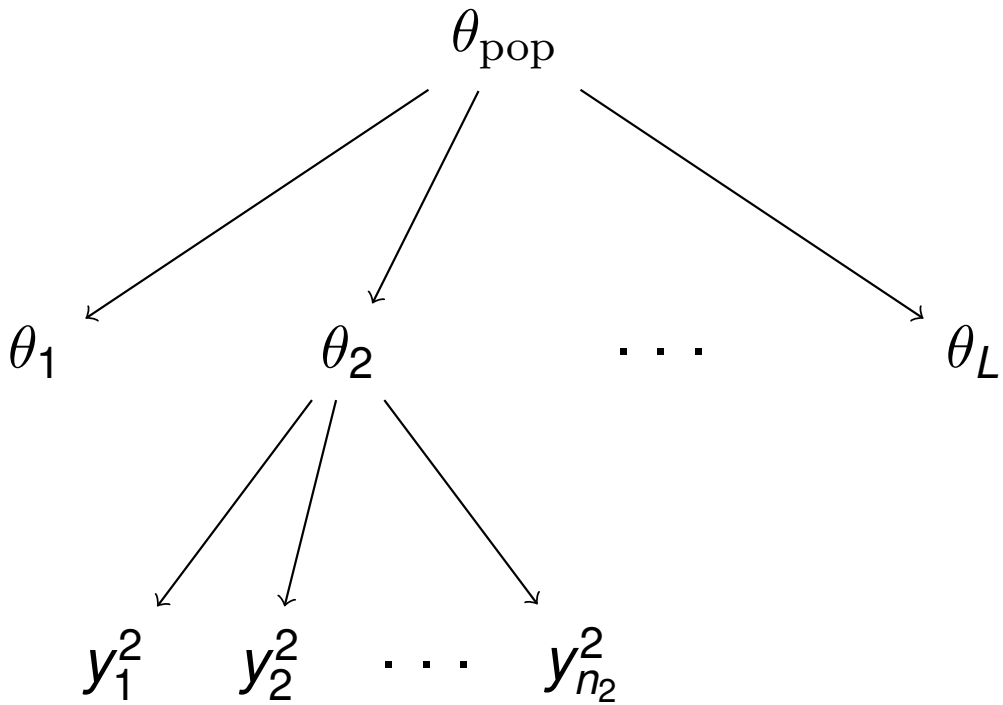
- do partial pooling.
- estimate how similar the groups are to one another.
- estimate individual parameters.

Hierarchical model

$$\theta = (\theta_1, ..., \theta_L) \sim p(\theta|\theta_{\text{pop}})$$

$$y = (y_1, ..., y_N) \sim p(y|\theta, x)$$

Hierarchical model



Example 3: Hierarchical two compartment model

Likelihood function:

$$\log\theta \sim \text{Normal}(\log\theta_{\text{pop}}, \Omega)$$

$$\Omega = \begin{pmatrix} \omega_1 & 0 & 0 & 0 & 0 \\ 0 & \omega_2 & 0 & 0 & 0 \\ 0 & 0 & \omega_3 & 0 & 0 \\ 0 & 0 & 0 & \omega_4 & 0 \\ 0 & 0 & 0 & 0 & \omega_5 \end{pmatrix}$$

$$\log(cObs) \sim \text{Normal}\left(\log\left(\frac{y_2}{VC}\right), \sigma^2\right)$$

Exercise 6: Write, fit, and diagnose a hierarchical two compartment model for a population of 10 patients. Use `data/twoCptPop.data.r` *and* `twoCptPop.r`.

- *Start by running 3 chains with 30 iterations.*
- *Do you get any warning messages?*

Divergent transitions

"There were 29 divergent transitions after warmup."

- A divergent transition occurs when we fail to accurately compute a Hamiltonian trajectory.
- This is because we *approximate* trajectories.
- Our sampler may not be refined enough to explore the entire typical set.

Divergent transitions

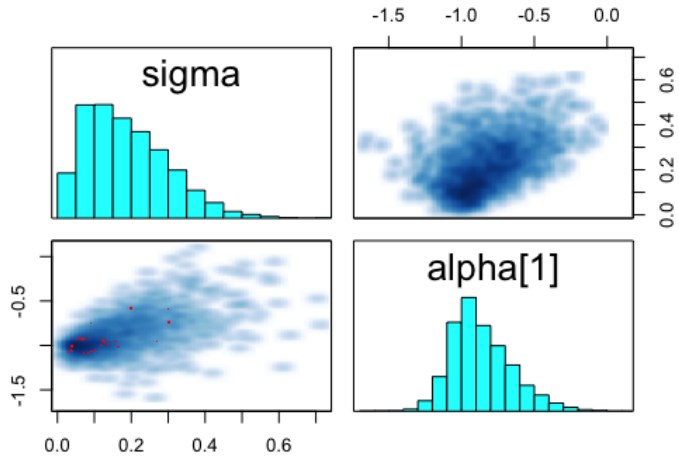
Consider the following hierarchical model:

$$\begin{aligned} \alpha_i &\sim \text{Normal}(\mu, \sigma) \\ y_i &\sim p(y|\alpha_i) \end{aligned}$$

Divergent transitons

$$\alpha_j \sim \text{Normal}(\mu, \sigma)$$

Fitting this model yields the following pairs plot:



- This geometric shape is known as Neil's funnel [Neil, 2003].
- Its interactions with HMC is described in [Betancourt and Girolmi, 2015].
- It occurs in hierarchical models when we have sparse data and a centered prior.

Proposition: reparametrize the model to avoid the funnel shape. We will do so by standardizing α.

$$\alpha_{\text{std},i} := \frac{\alpha_i - \mu}{\sigma}$$

Then

$$\alpha_{\text{std}} \sim \text{Normal}(0, 1)$$

Then

$$\alpha_i = \mu + \sigma\alpha_{\mathrm{std},i}$$

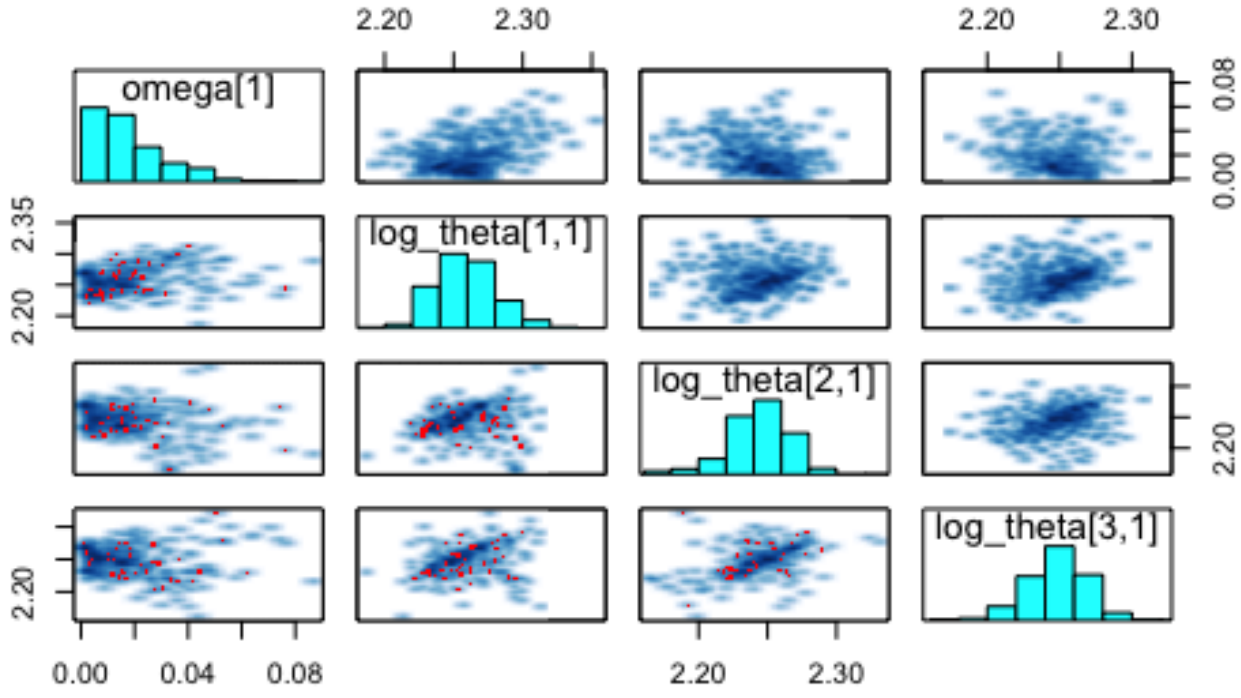
Hence

$$y_i \sim p(\mu + \sigma\alpha_{\mathrm{std},i})$$

- Same data generating process; but how does this affect the geometry of the posterior?

Our model is a little more complicated than the above example:

- a lot of parameters (100 +)!
- multiple population parameters and hierarchical structures.
- these parameters follow a log normal distribution (so we need a pairs plot with $\log \theta$).


omega[1]
log_theta[1,1]
log_theta[2,1]
log_theta[3,1]
2.20
2.30
0.00
0.08
2.35
2.20
0.04

Exercise 6: Reparametrize the two compartment population model and fit it.

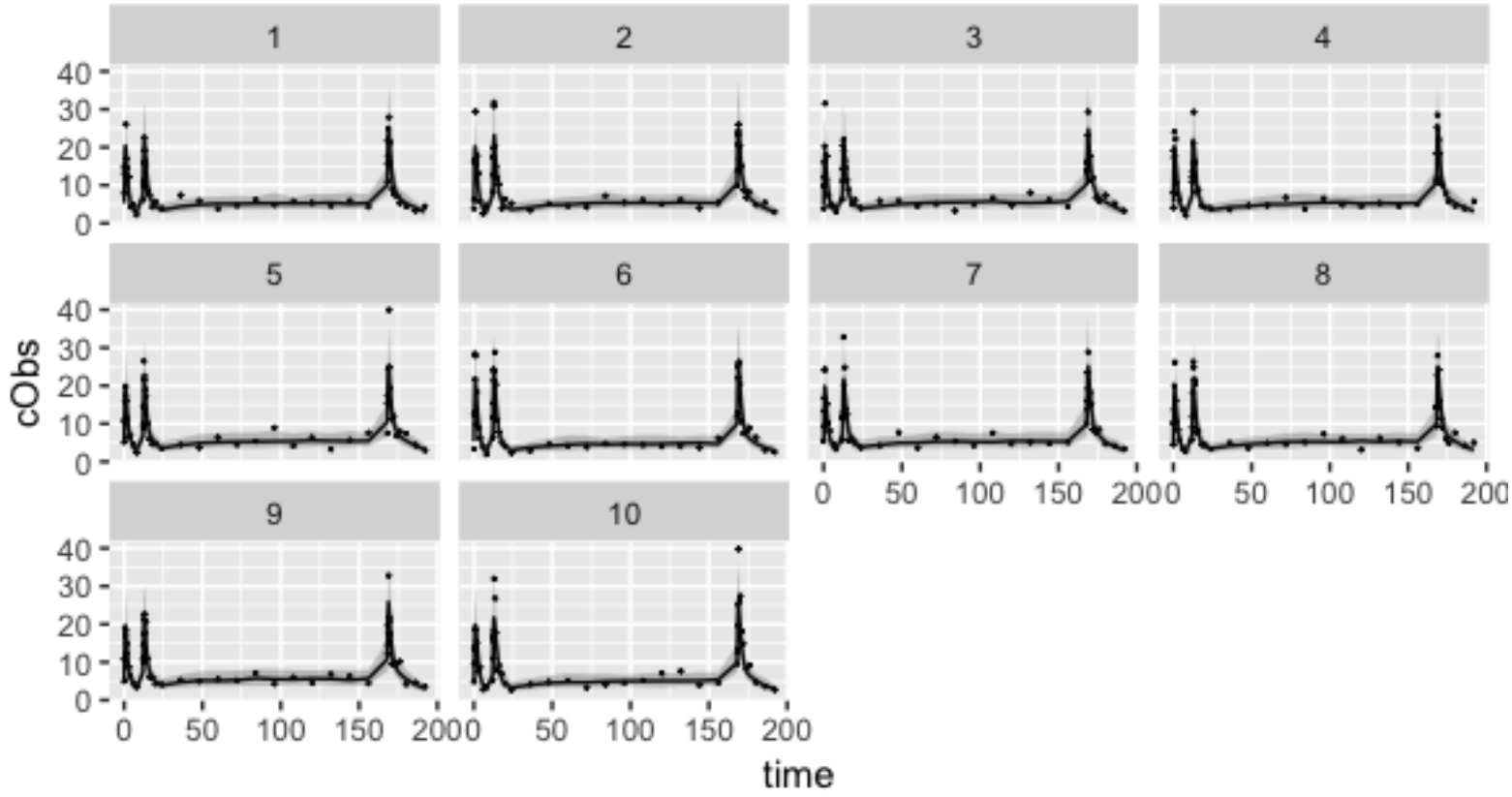
- First, work out the appropriate parametrization. You should start with $\log \theta_i \sim \mathrm{Normal}(\theta_{\mathrm{pop},i}, \omega)$
- Write, fit, and check the inference (run 100 chains).
- What kind of predictive checks can we do?

Need:

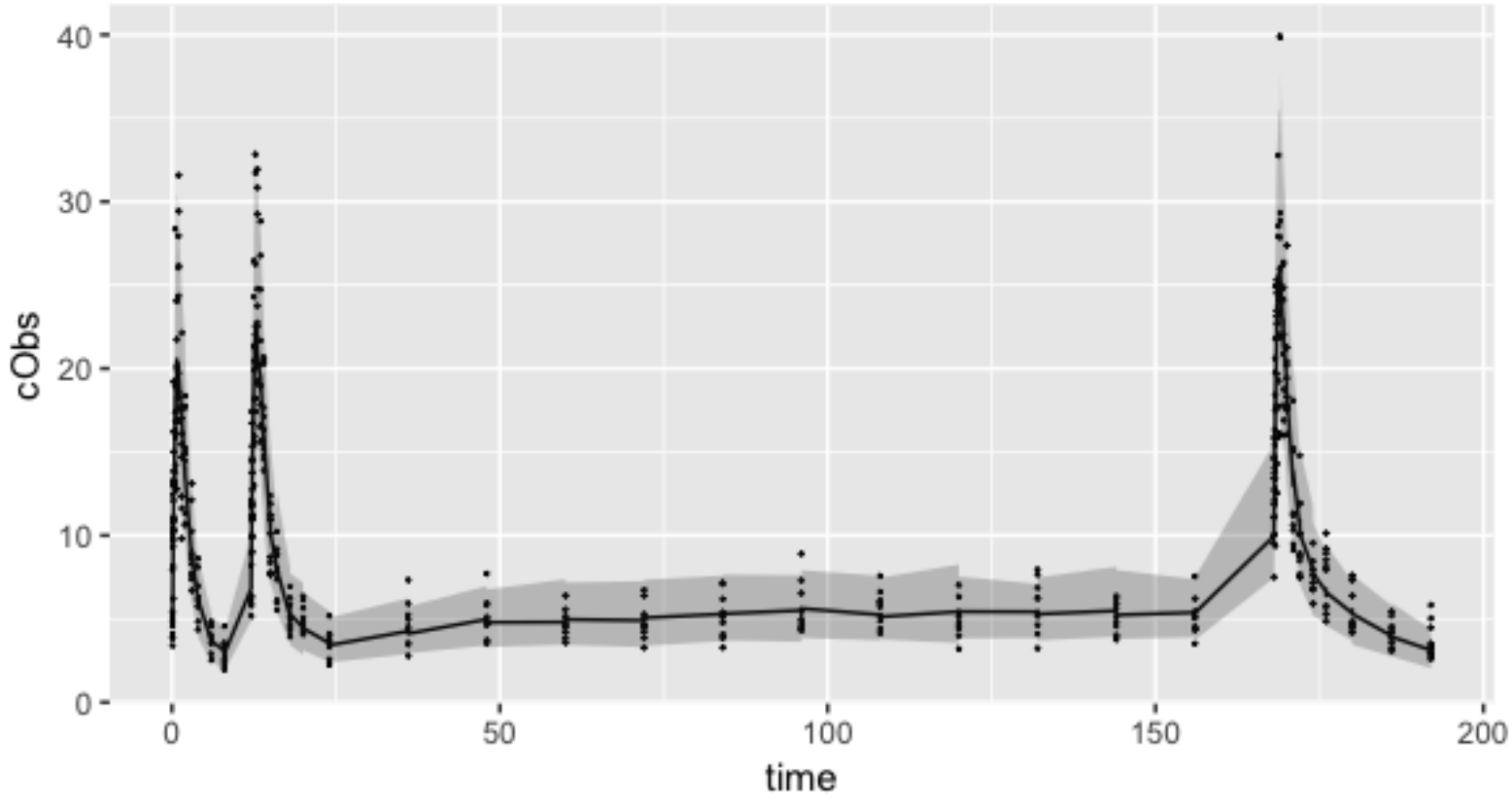
- predictions at an individual level
- predictions at a population level

As always, this comes down to properly writing the data generating process in the generated quantities block.

Individual predictions



Population predictions



- For a very good case study on hierarchical models, see, Bob Carpenter's *Pooling with Hierarchical Models for Repeated Binary Trials*
- https://mc-stan.org/users/documentation/case-studies/pool-binary-trials.html

References I

Betancourt, M. and Girolmi, M. (2015).
Hamiltonian monte carlo for hierarchical models.
Current trends in Bayesian methodology with applications, 79.

Neil, R. M. (2003).
Slice sampling.
Annals of Statistics, 31.