Stan by Example

Structure of Session

- Participate
  - try
  - ask questions
- I type, you type
- Exercises as we go
- R, Stan files available online to download after

Setup

- Start RStudio
- Load RStan: `library(rstan)`

Stan Process

- In Stan language, specify joint dist $p(\theta, y)$
- Pass to Stan data, model, get correlated samples $p(\theta \mid y)$

Bernoulli example

Making Inferences

- $\mathbb{E}[g(\theta) \mid y] \approx \frac{1}{M} \sum_{m=1}^{M} g(\theta^{(m)})$
- Posterior mean
- Posterior quantiles
- Generated quantities

What is the posterior probability of getting 5 ones?, $\Pr[\sum_{i}^{10} \tilde{y}_i = 5 \mid y]$

Priors

- Default prior is **flat** - no contribution to log probability
- Priors go in model block
- With small amounts of data, prior has large impact

Compare posterior of θ under flat prior with beta$(5, 5)$

What is RStan doing?

- Read Stan program
- Translate to C++
- Compile C++, link to R
- Run sampler with data, return values