Hierarchical Models

PSYC 573

University of Southern California
March 3, 2022

Therapeutic Touch Example (N = 28)

Data Points From One Person

y: whether the guess of which hand was hovered over was correct

Person S01

y	s
1	S01
0	S01
0	S01
0	S01
0	S01
0	S01
0	S01
0	S01
0	S01
0	S01

Binomial Model

We can use a Bernoulli model:

$$y_i \sim \text{Bern}(\theta)$$

for $i = 1, \ldots, N$

Assuming exchangeability given θ, more succint to write

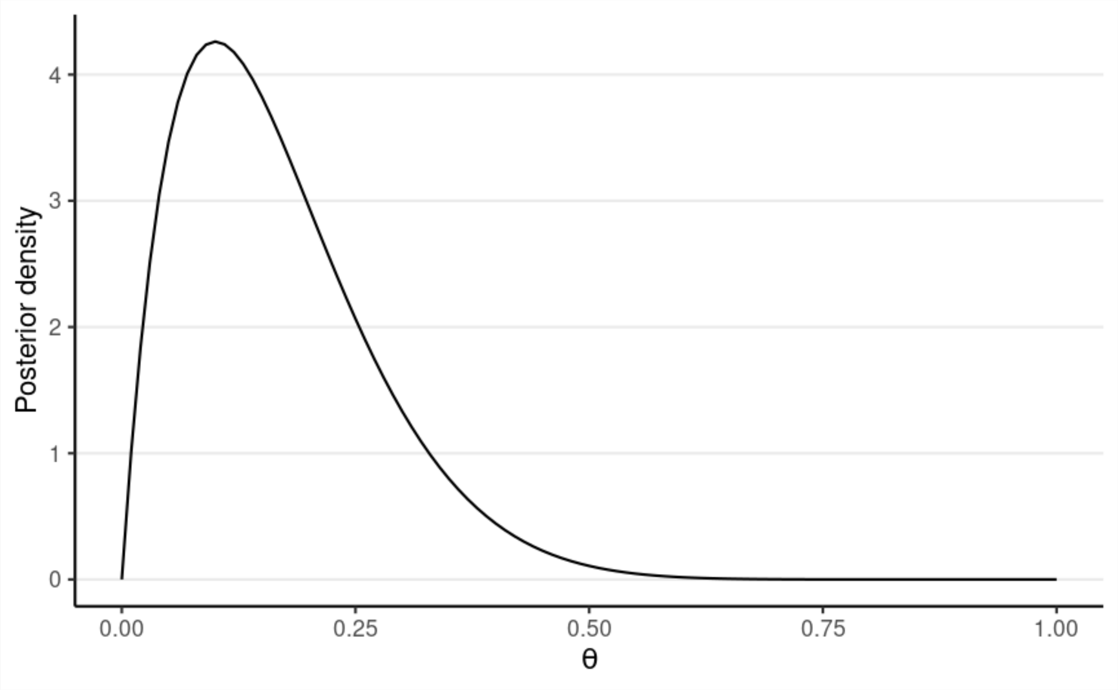
$$z \sim \text{Bin}(N, \theta)$$

for $z = \sum_{i=1}^{N} y_i$

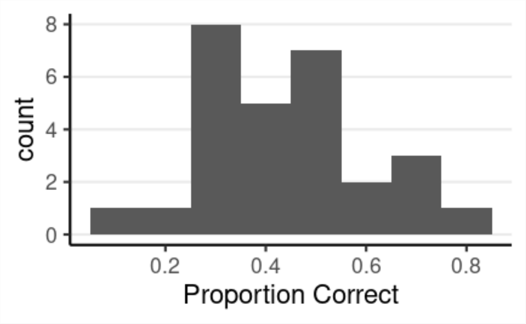
- Bernoulli: Individual trial
- Binomial: total count of "1"s

1 success, 9 failures

Posterior: Beta(2, 10)



Multiple People



We could repeat the binomial model for each of the 28 participants, to obtain posteriors for $\theta_1, \ldots, \theta_{28}$

But . . .

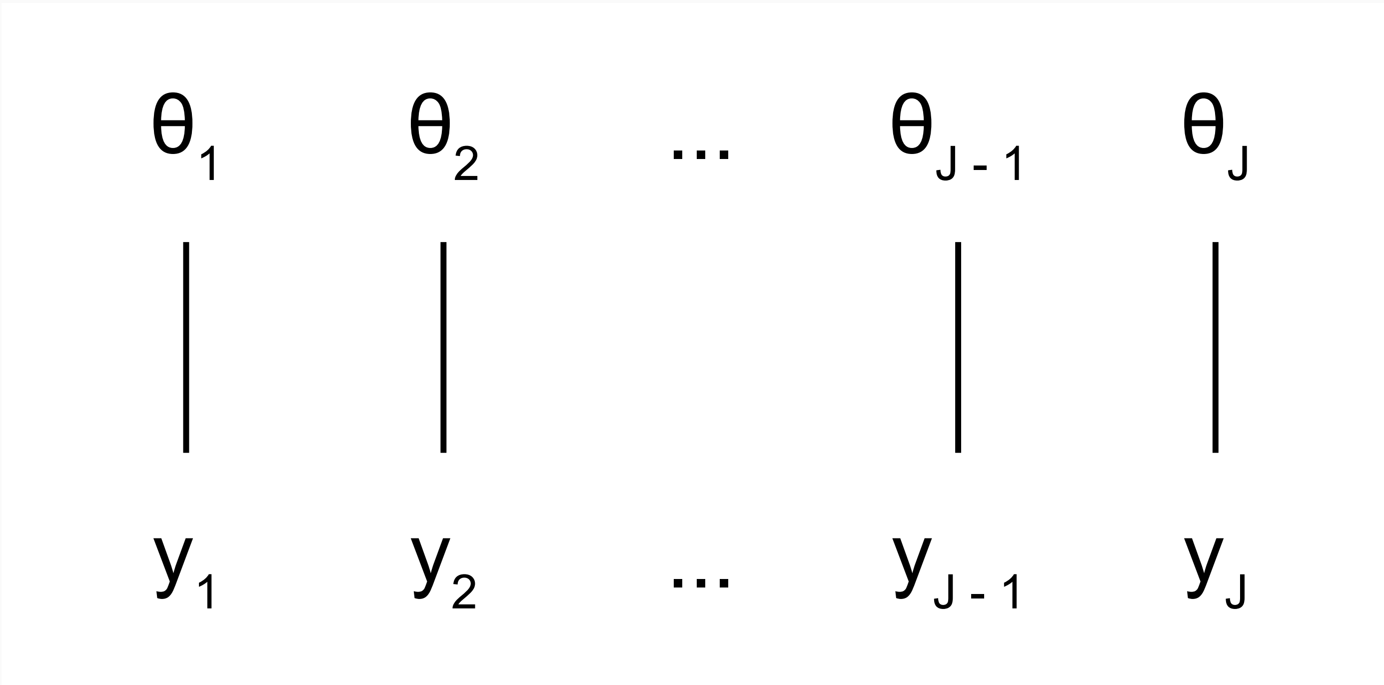
> Do we think our belief about θ_1 would inform our belief about θ_2, etc?

After all, human beings share 99.9% of genetic makeup

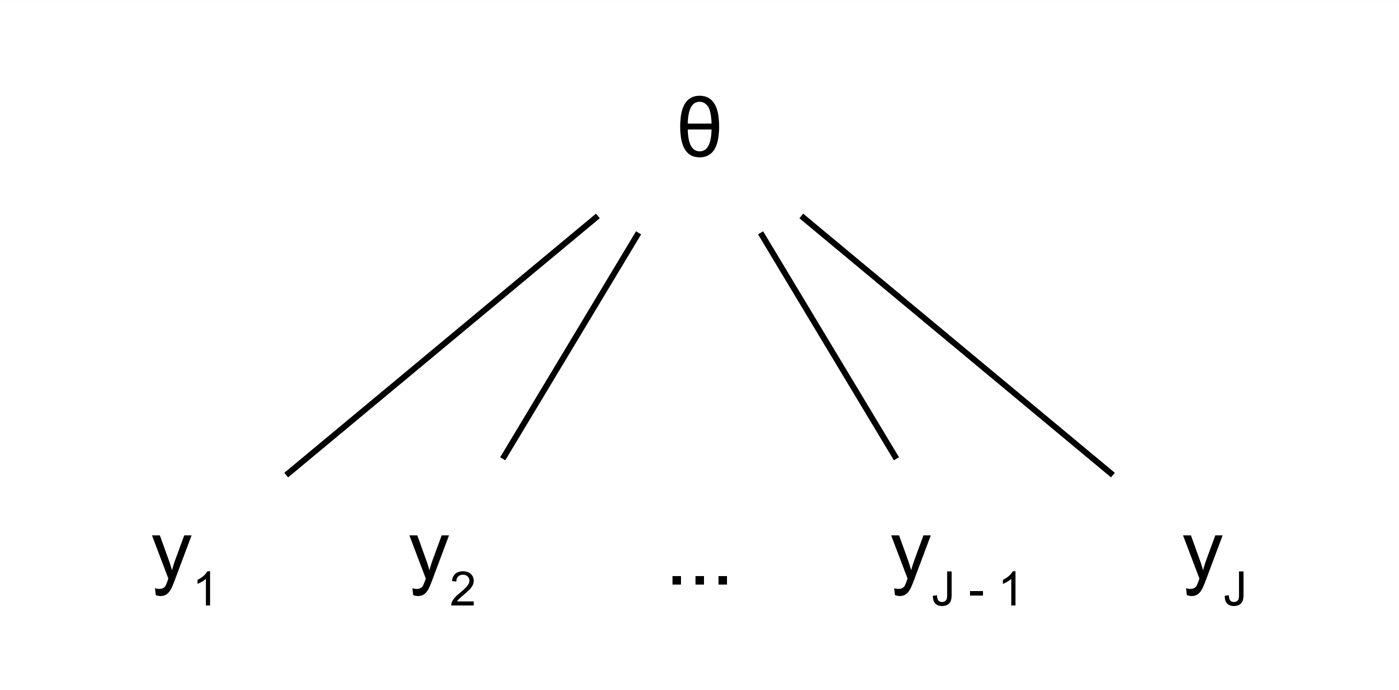
Three Positions of Pooling

- No pooling: each individual is completely different; inference of θ_1 should be independent of θ_2, etc
- Complete pooling: each individual is exactly the same; just one θ instead of 28 θ_j's
- **Partial pooling**: each individual has something in common but also is somewhat different

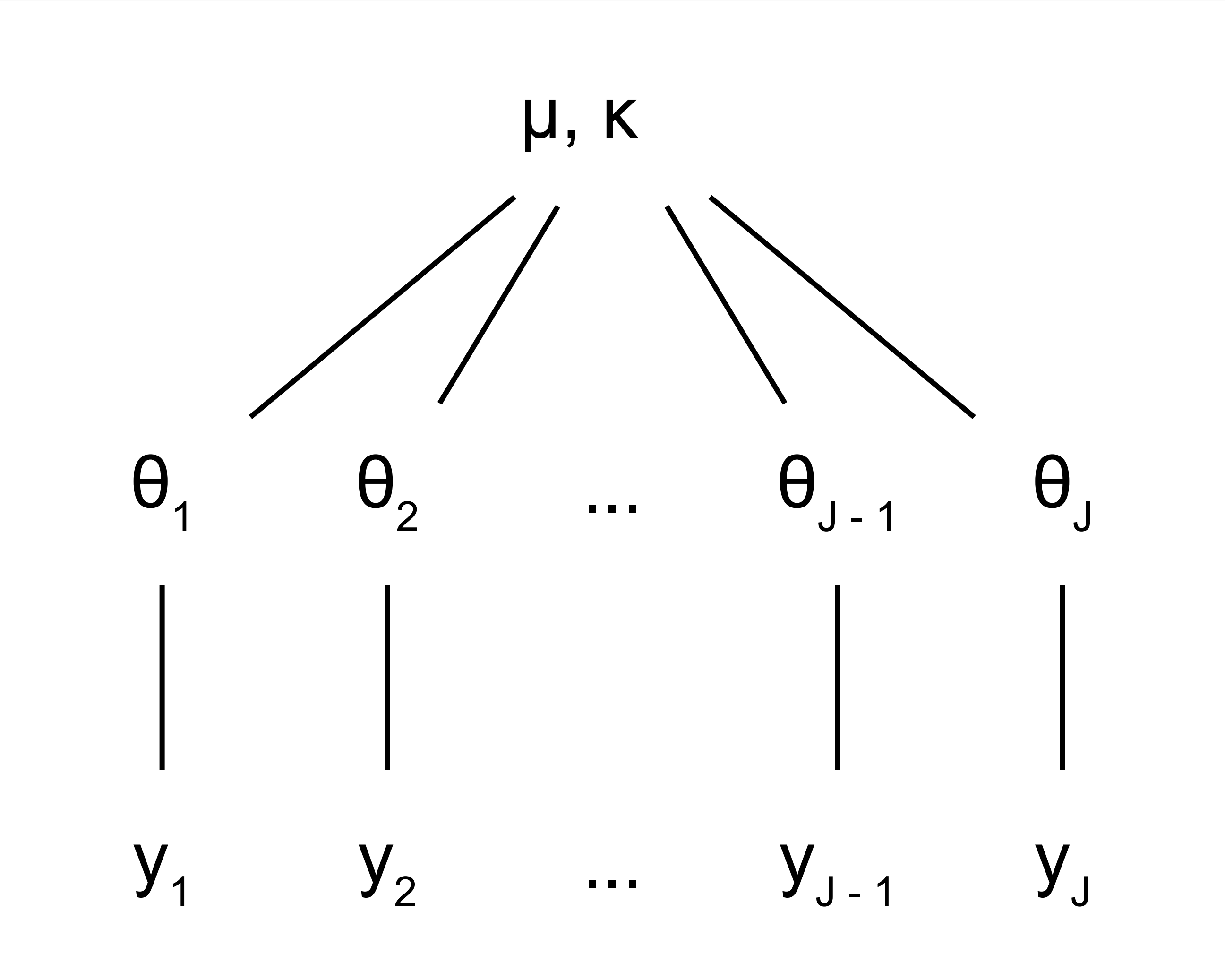
No Pooling



Complete Pooling



Partial Pooling



Partial Pooling in Hierarchical Models

Hierarchical Priors: $\theta_j \sim \mathbf{Beta2}(\mu, \kappa)$

Beta2: *reparameterized* Beta distribution

- mean $\mu = a/(a+b)$
- concentration $\kappa = a+b$

Expresses the prior belief:

> Individual θs follow a common Beta distribution with mean μ and concentration κ

How to Choose κ

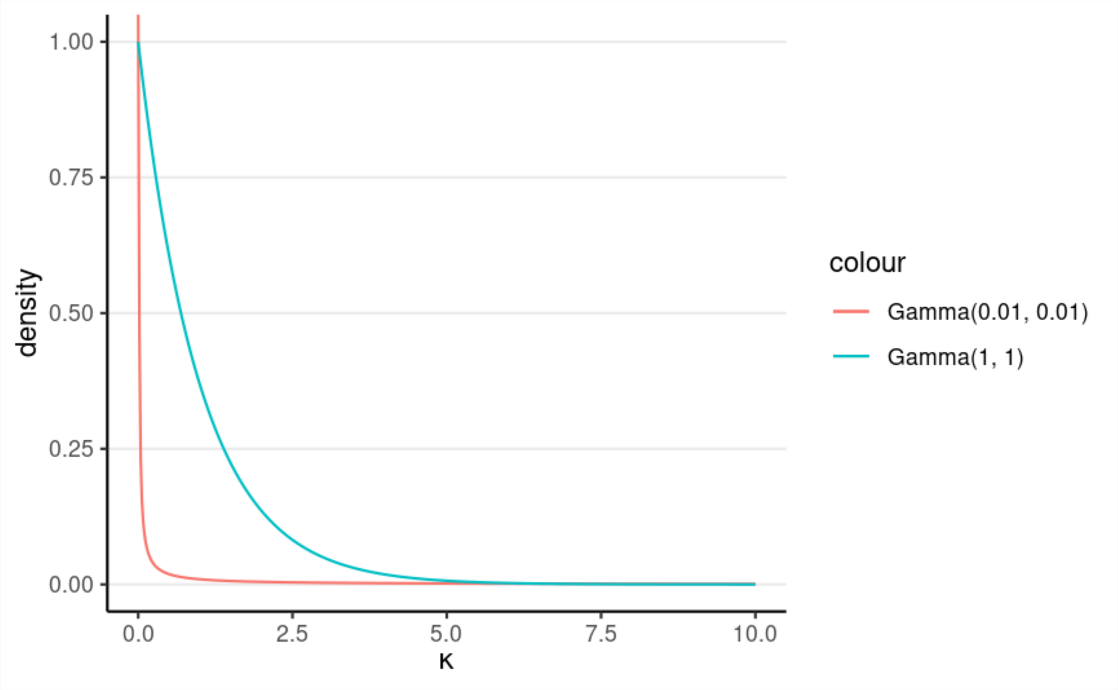
If $\kappa \to \infty$: everyone is the same; no individual differences (i.e., complete pooling)

If $\kappa = 0$: everybody is different; nothing is shared (i.e., no pooling)

We can fix a κ value based on our belief of how individuals are similar or different

A more Bayesian approach is to treat κ as an unknown, and use Bayesian inference to update our belief about κ

Generic prior by Kruschke (2015): $\kappa \sim$ Gamma(0.01, 0.01)



Sometimes you may want a stronger prior like Gamma(1, 1), if it is unrealistic to do no pooling

Full Model

Model   Stan code

Model:

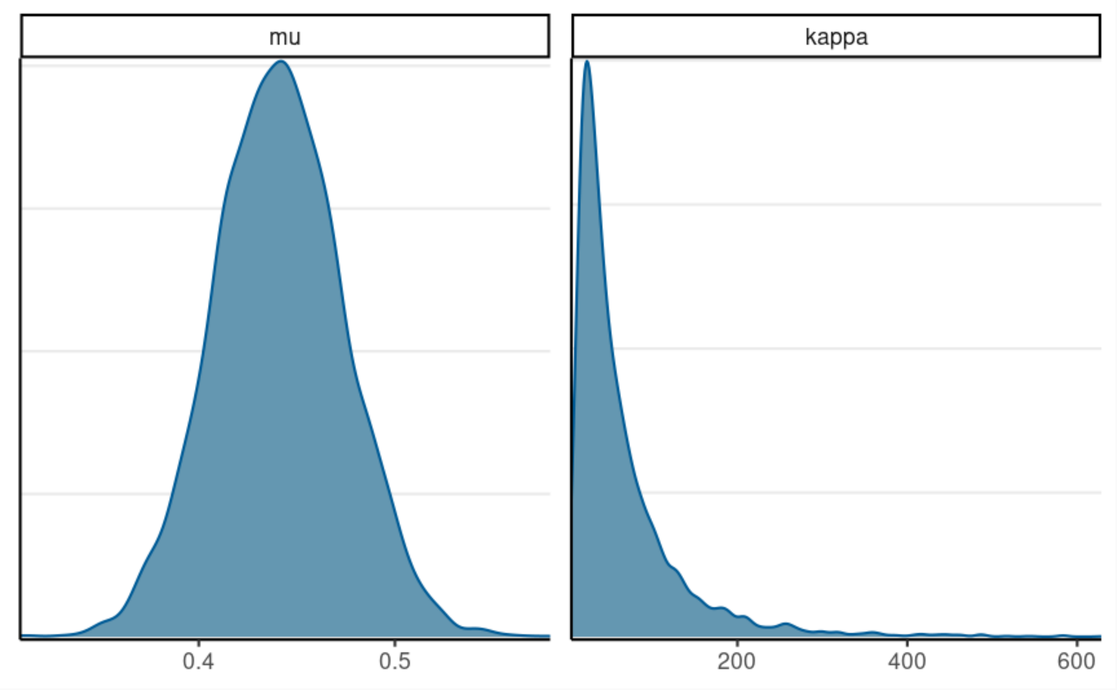
$$z_j \sim \mathrm{Bin}(N_j, \theta_j)$$
$$\theta_j \sim \mathrm{Beta2}(\mu, \kappa)$$

Prior:

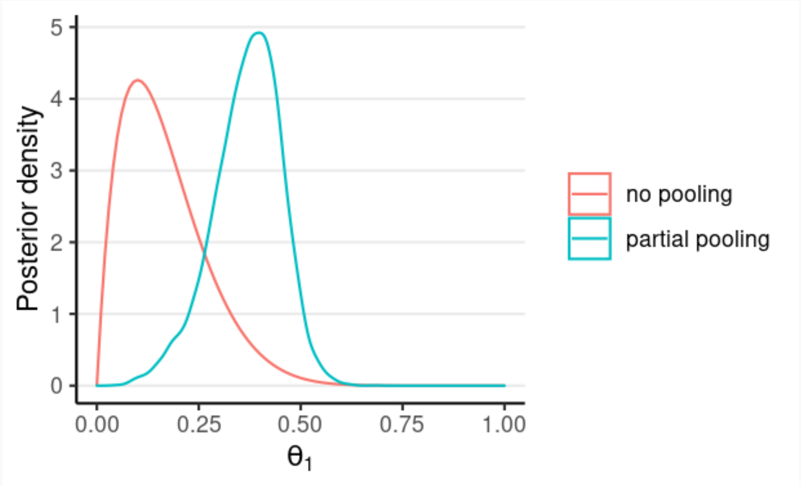
$$\mu \sim \mathrm{Beta}(1.5, 1.5)$$
$$\kappa \sim \mathrm{Gamma}(0.01, 0.01)$$

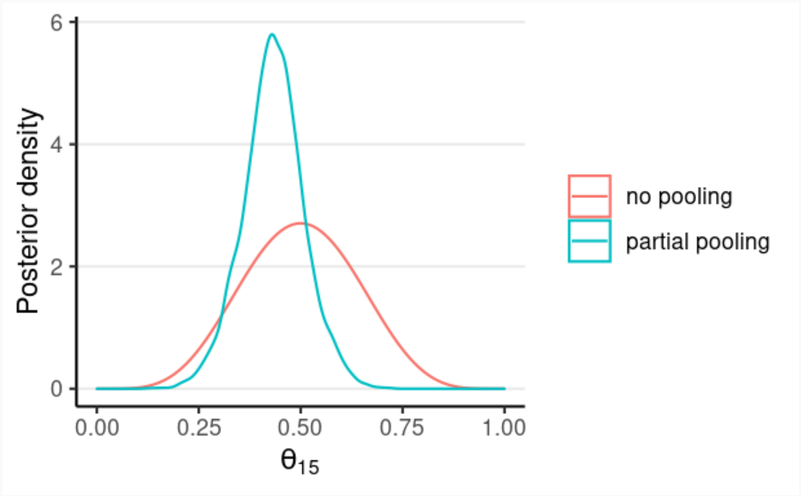
Posterior of Hyperparameters

```r
library(bayesplot)
mcmc_dens(tt_fit, pars = c("mu", "kappa"))
```



Shrinkage


Posterior density
5
4
3
2
1
0
0.00
0.25
0.50
0.75
1.00
θ1
no pooling
partial pooling


Posterior density
6
4
2
0
0.00
0.25
0.50
0.75
1.00
θ15
no pooling
partial pooling

Multiple Comparisons?

Frequentist: family-wise error rate depends on the number of intended contrasts

Bayesian: only one posterior; hierarchical priors already express the possibility that groups are the same

> Thus, Bayesian hierarchical model "completely solves the multiple comparisons problem."[1]

[1]: see https://statmodeling.stat.columbia.edu/2016/08/22/bayesian-inference-completely-solves-the-multiple-comparisons-problem/

[2]: See more in ch 11.4 of Kruschke (2015)

Hierarchical Normal Model

Effect of coaching on SAT-V

School	Treatment Effect Estimate	Standard Error
A	28	15
B	8	10
C	-3	16
D	7	11
E	-1	9
F	1	11
G	18	10
H	12	18

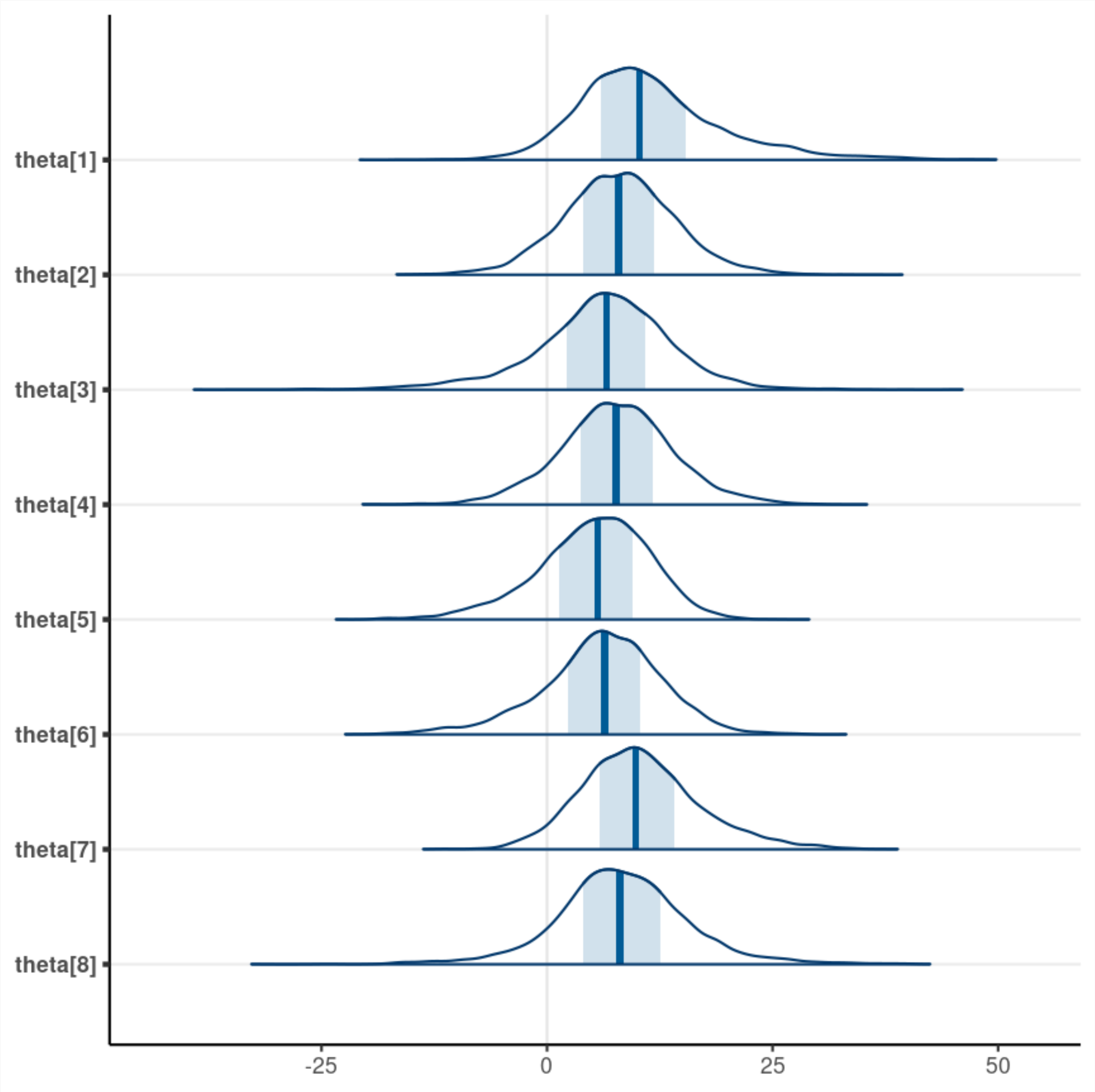
Model   Stan code

Model:

$$
\begin{aligned}
d_j &\sim N(\theta_j, s_j) \\
\theta_j &\sim N(\mu, \tau)
\end{aligned}
$$

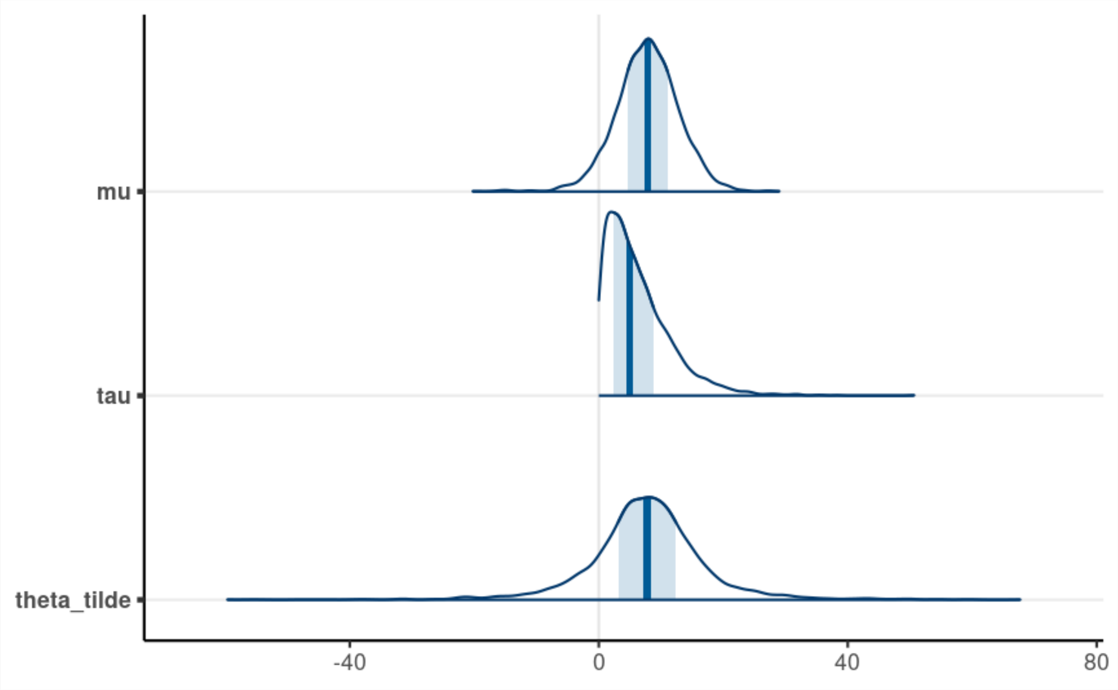
Prior:

$$
\begin{aligned}
\mu &\sim N(0, 100) \\
\tau &\sim t_4^+(0, 100)
\end{aligned}
$$


theta[1]
theta[2]
theta[3]
theta[4]
theta[5]
theta[6]
theta[7]
theta[8]
-25
0
25
50

Prediction Interval

Posterior distribution of the true effect size of a new study, $\tilde{\theta}$



See https://onlinelibrary.wiley.com/doi/abs/10.1002/jrsm.12 for an introductory paper on random-effect meta-analysis