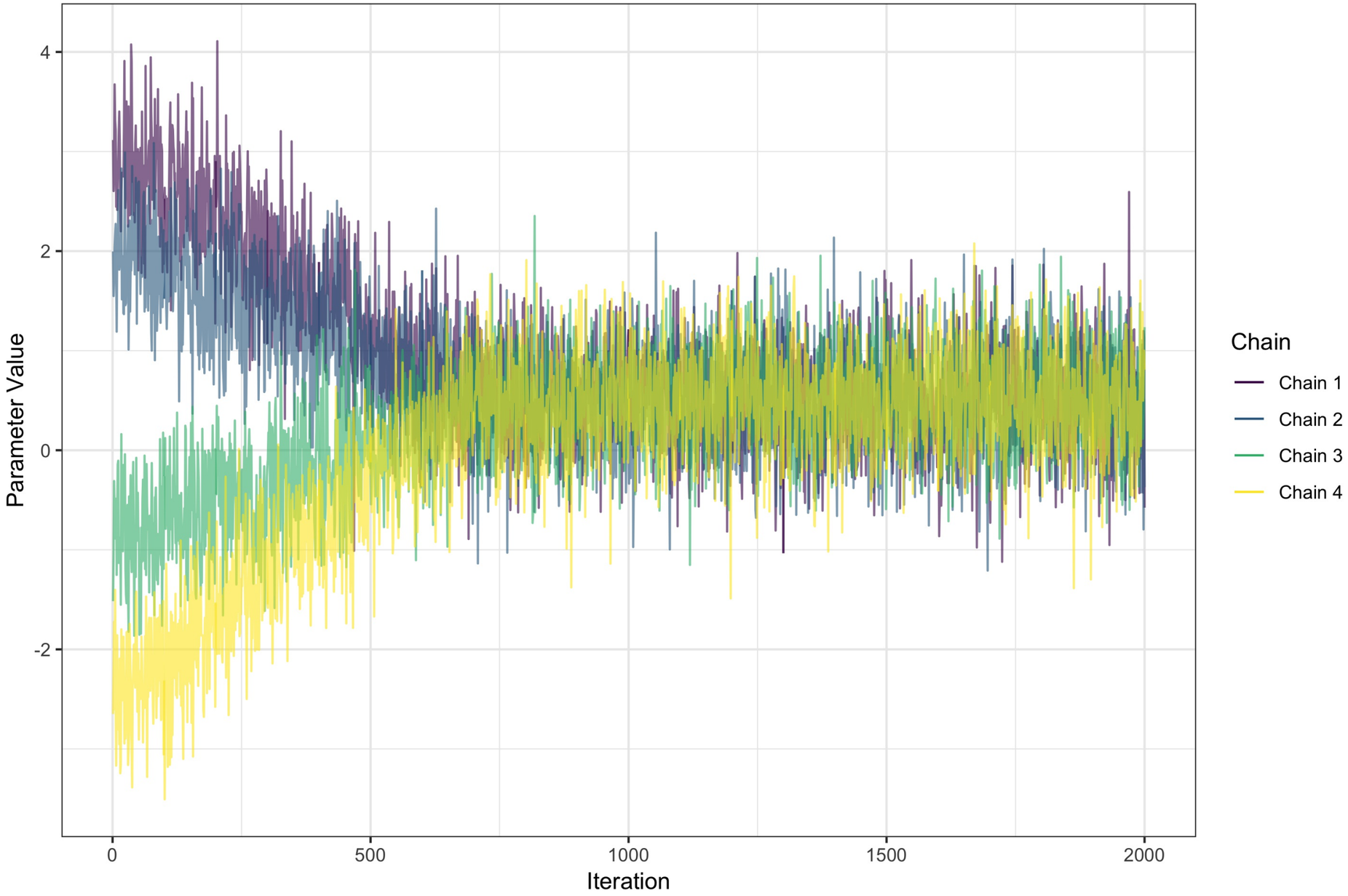
BAYESIAN REGRESSION MODELING WITH RSTANARM

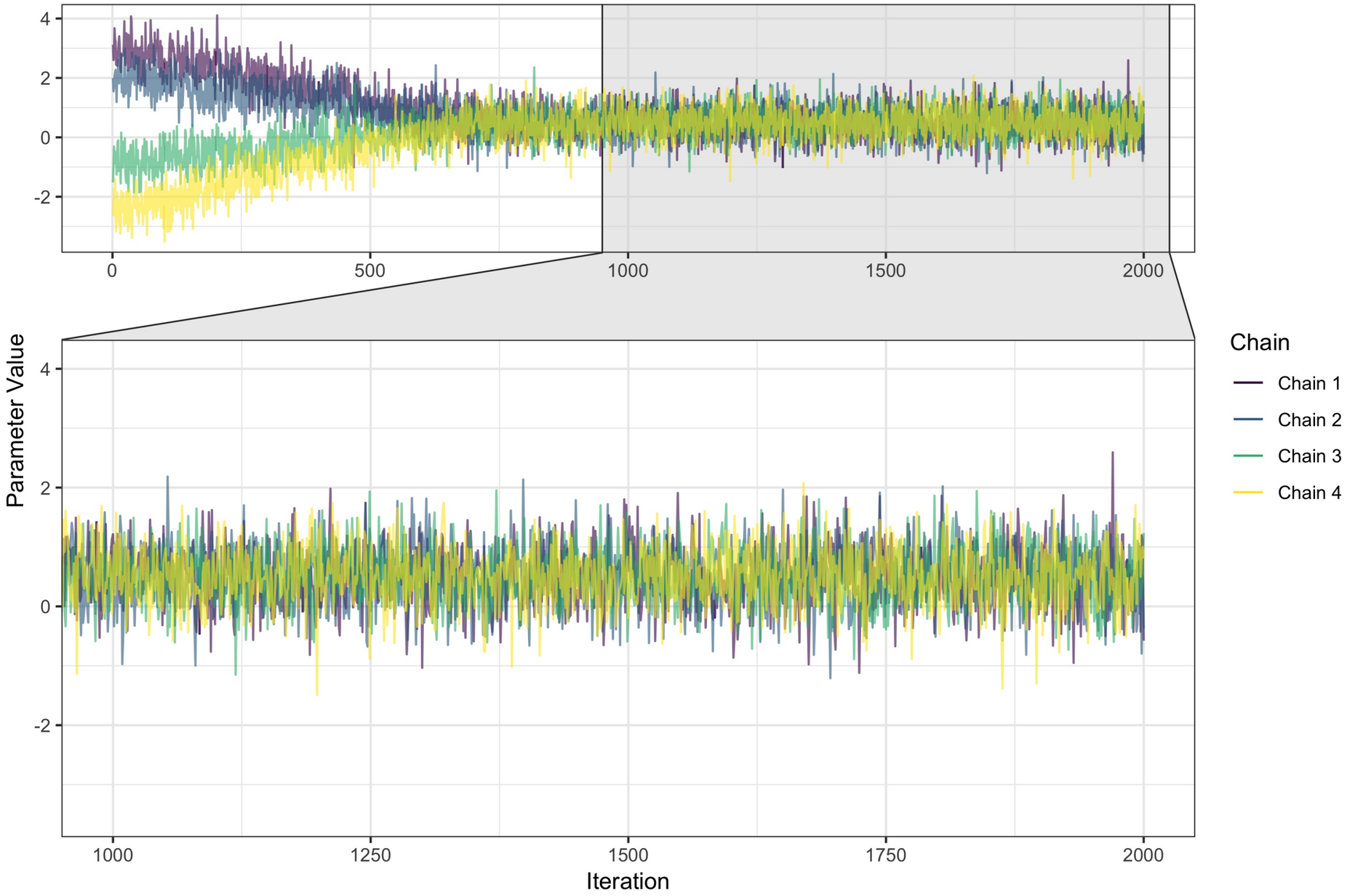
What's in a Bayesian Model?

Jake Thompson
Psychometrician, ATLAS, University of Kansas

Posterior Distributions

- Posterior distributions sampled in groups called chains
- Each sample in a chain is an iteration


Parameter Value
4
2
0
-2
0
500
1000
1500
2000
Iteration
Chain
Chain 1
Chain 2
Chain 3
Chain 4

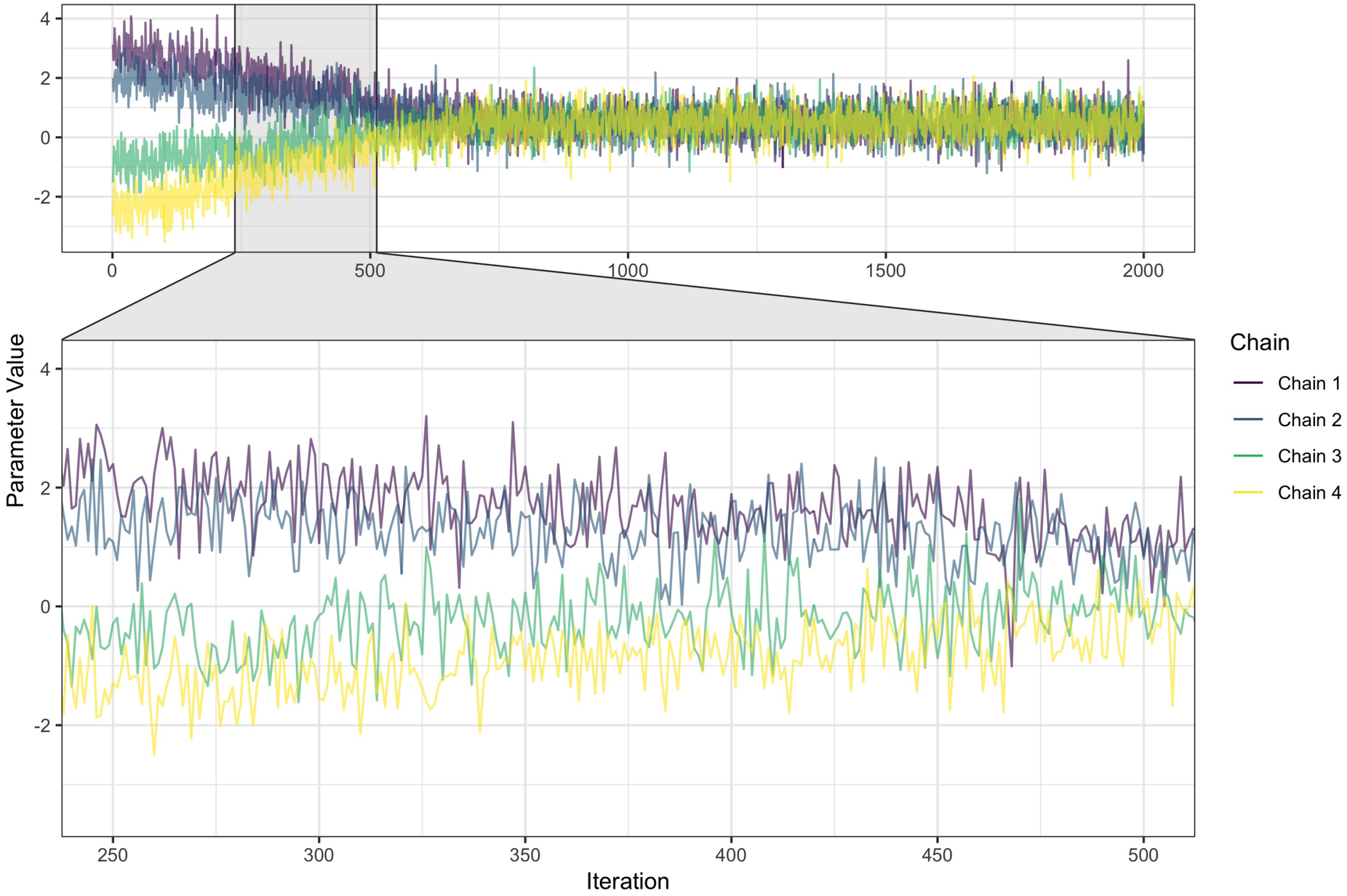

4
2
0
-2
0
500
1000
1500
2000
Parameter Value
4
2
0
-2
1000
1250
1500
1750
2000
Iteration
Chain
Chain 1
Chain 2
Chain 3
Chain 4

Changing the Number and Length of Chains

```
stan_model <- stan_glm(kid_score ~ mom_iq, data = kidiq,
  chains = 3, iter = 1000, warmup = 500)
```

Changing the Number and Length of Chains

```
summary(stan_model)
#> Model Info:
#>
#>  function:     stan_glm
#>  family:       gaussian [identity]
#>  formula:      kid_score ~ mom_iq
#>  algorithm:    sampling
#>  priors:       see help('prior_summary')
#>  sample:       1500 (posterior sample size)
#>  observations: 434
#>  predictors:   2
#>
#> Estimates:
#>                mean    sd      2.5%    25%     50%     75%     97.5%
#> (Intercept)     25.8     6.0    14.1    21.7    25.6    29.9    37.5
#> mom_iq           0.6     0.1     0.5     0.6     0.6     0.7     0.7
#> sigma           18.3     0.6    17.2    17.9    18.3    18.7    19.6
#> mean_PPD        86.9     1.3    84.5    86.0    86.9    87.7    89.2
#> log-posterior -1885.4     1.2 -1888.4 -1885.9 -1885.1 -1884.5 -1884.0
#>
#> Diagnostics:
#>               mcse Rhat n_eff
#> (Intercept)   0.2  1.0  1500
#> mom_iq        0.0  1.0  1500
#> sigma         0.0  1.0  1500
```


4
2
0
-2
0
500
1000
1500
2000
Parameter Value
250
300
350
400
450
500
Iteration
Chain
Chain 1
Chain 2
Chain 3
Chain 4

How many iterations?

- Fewer iterations = shorter estimation time
- Not enough iteration = convergence problems

BAYESIAN REGRESSION MODELING WITH RSTANARM

Let's practice!

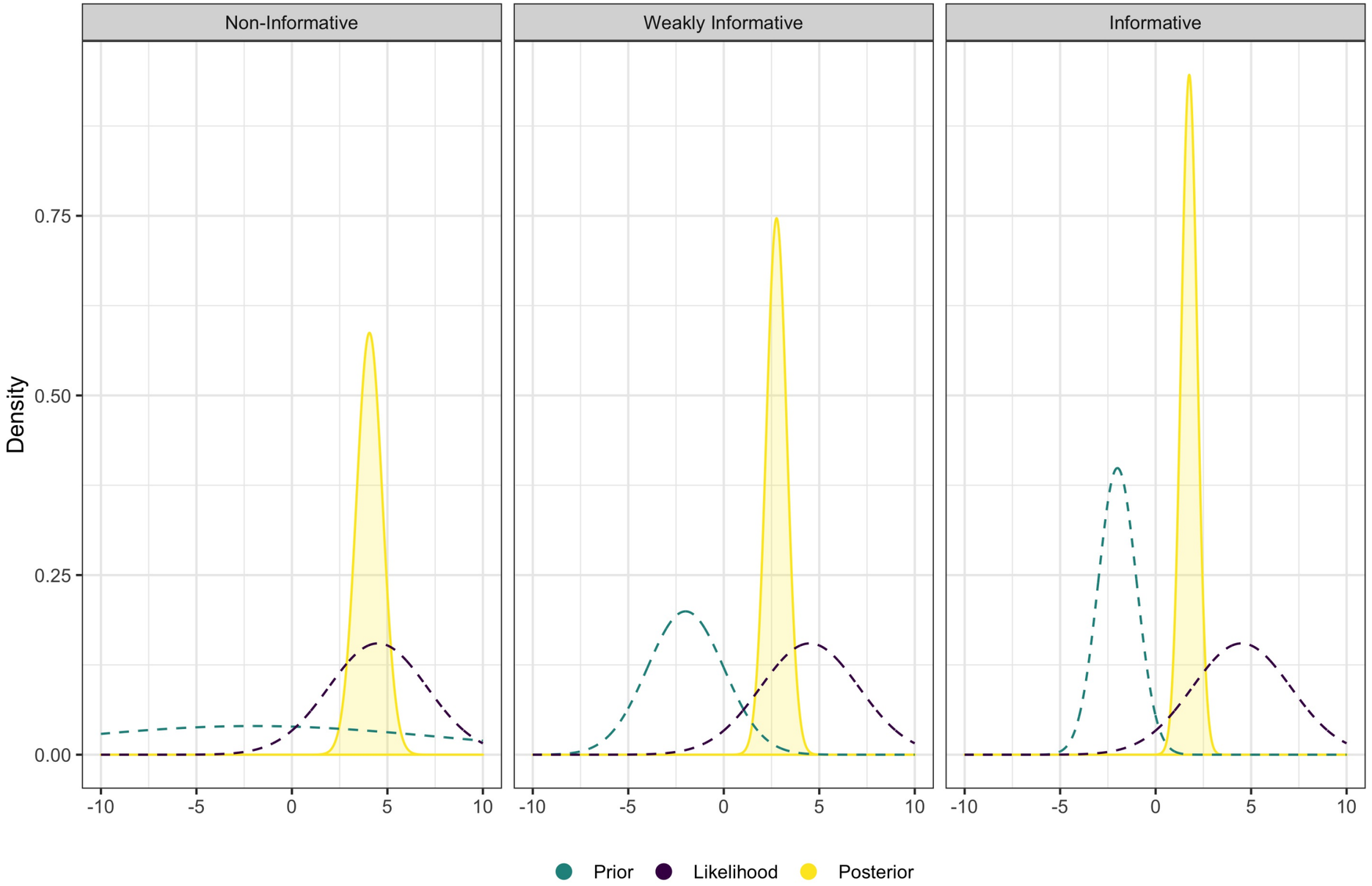
BAYESIAN REGRESSION MODELING WITH RSTANARM

Prior Distributions

Jake Thompson
Psychometrician, ATLAS, University of Kansas

What's a prior distribution?

- Information that we bring to the model
- Likelihood + prior = posterior


Non-Informative
Weakly Informative
Informative
Density
0.75
0.50
0.25
0.00
-10
-5
0
5
10
Prior
Likelihood
Posterior

Prior Distrubtions in rstanarm

```
stan_model <- stan_glm(kid_score ~ mom_iq, data = kidiq)

prior_summary(stan_model){{1}}
#> Priors for model 'stan_model'
#> ------
#> Intercept (after predictors centered)
#>  ~ normal(location = 0, scale = 10)
#>      **adjusted scale = 204.11
#>
#> Coefficients
#>  ~ normal(location = 0, scale = 2.5)
#>      **adjusted scale = 3.40
#>
#> Auxiliary (sigma)
#>  ~ exponential(rate = 1)
#>      **adjusted scale = 20.41 (adjusted rate = 1/adjusted scale)
#> ------
#> See help('prior_summary.stanreg') for more details
```

Calculating Adjusted Scales

- Intercept: 10 * sd(y)
- Coefficients: (2.5 / sd(x)) * sd(y)

```
prior_summary(stan_model)
#> Priors for model 'stan_model'
#> ------
#> Intercept (after predictors centered)
#>  ~ normal(location = 0, scale = 10)
#>      **adjusted scale = 204.11
#>
#> Coefficients
#>  ~ normal(location = 0, scale = 2.5)
#>      **adjusted scale = 3.40
```

```
10 * sd(kidiq$kid_score)
#> [1] 204.1069
```

```
(2.5 / sd(kidiq$mom_iq)) * sd(kidiq$kid_score)
#> [1] 3.401781
```

Unadjusted Priors

```
no_scale <- stan_glm(kid_score ~ mom_iq, data = kidiq,
  prior_intercept = normal(autoscale = FALSE),
  prior = normal(autoscale = FALSE),
  prior_aux = exponential(autoscale = FALSE)
)
```

```
prior_summary(no_scale)
#> Priors for model 'no_scale'
#> ------
#> Intercept (after predictors centered)
#>  ~ normal(location = 0, scale = 10)
#>
#> Coefficients
#>  ~ normal(location = 0, scale = 2.5)
#>
#> Auxiliary (sigma)
#>  ~ exponential(rate = 1)
#> ------
#> See help('prior_summary.stanreg') for more details
```

BAYESIAN REGRESSION MODELING WITH RSTANARM

Let's practice!

BAYESIAN REGRESSION MODELING WITH RSTANARM

User Specified Priors

Jake Thompson
Psychometrician, ATLAS, University of Kansas

Why change the default prior?

- Good reason to believe the parameter will take a given value
- Contraints on parameter

Specify a prior

```
stan_model <- stan_glm(kid_score ~ mom_iq, data = kidiq,
  prior_intercept = normal(location = 0, scale = 10),
  prior = normal(location = 0, scale = 2.5),
  prior_aux = exponential(rate = 1)
)
```

Specify a prior

```
stan_model <- stan_glm(kid_score ~ mom_iq, data = kidiq,
  prior_intercept = normal(location = 0, scale = 10, autoscale = FALSE),
  prior = normal(location = 0, scale = 2.5, autoscale = FALSE),
  prior_aux = exponential(rate = 1, autoscale = FALSE)
)
```

Specify a prior

```
stan_model <- stan_glm(kid_score ~ mom_iq, data = kidiq,
  prior_intercept = normal(location = 3, scale = 2),
  prior = cauchy(location = 0, scale = 1),
)
```

- Many different priors
  - normal()
  - exponential()
  - student_t()
  - cauchy()
- ?priors

Flat priors

```
stan_model <- stan_glm(kid_score ~ mom_iq, data = kidiq,
  prior_intercept = NULL,
  prior = NULL,
  prior_aux = NULL
)
```

```
prior_summary(stan_model)
#> Priors for model 'stan_model'
#> ------
#> Intercept (after predictors centered)
#>  ~ flat
#>
#> Coefficients
#>  ~ flat
#>
#> Auxiliary (sigma)
#>  ~ flat
#> ------
#> See help('prior_summary.stanreg') for more details
```

BAYESIAN REGRESSION MODELING WITH RSTANARM

Let's practice!

BAYESIAN REGRESSION MODELING WITH RSTANARM

Altering the Estimation Process

Jake Thompson
Psychometrician, ATLAS, University of Kansas

Divergent Transitions

```
1: There were 15 divergent transitions after warmup. Increasing adapt_delta
above 0.8 may help.
```

- Too big of steps in the estimator
- Adjust step size

```
stan_model <- stan_glm(popularity ~ song_age, data = songs,
  control = list(adapt_delta = 0.95))
```

```
stan_model <- stan_glm(popularity ~ song_age, data = songs,
  control = list(adapt_delta = 0.99))
```

Exceeding the Maximum Treedepth

```
Chain 1 reached the maximum tree depth
```

- Sample evaluates branches and looks for a good place to "U-Turn"
- Max tree depth indicates poor efficiency

```
stan_model <- stan_glm(popularity ~ song_age, data = songs,
  control = list(max_treedepth = 10))
```

```
stan_model <- stan_glm(popularity ~ song_age, data = songs,
  control = list(max_treedepth = 15))
```

Tuning the Estimation

- Estimation errors are threats to the validity of the model
- Although complicated, these errors can be addressed easily

BAYESIAN REGRESSION MODELING WITH RSTANARM

Let's practice!