



A Bayesian Approach to Stats

Aj Wright

Goals

- Main differences between frequentist and Bayesian approaches
- How to specify a model in brms
- How to interpret brms output
- Pros vs cons of Bayesian approaches

Recap of Frequentist Stats

- Some assumptions
- Need large sample size to benefit from central limit theorem
- Uniform probability of parameter values (from the start, all values have an equal chance of being the parameter value)
- Ordinary Least Squares (OLS) estimation (or Maximum Likelihood estimation)
- Point estimates with associated uncertainty of this estimate
- p-values and confidence intervals
- lm /glm (or lme4/nlme/glmer for multilevel/mixed effects modeling)

Basics of Bayesian Stats

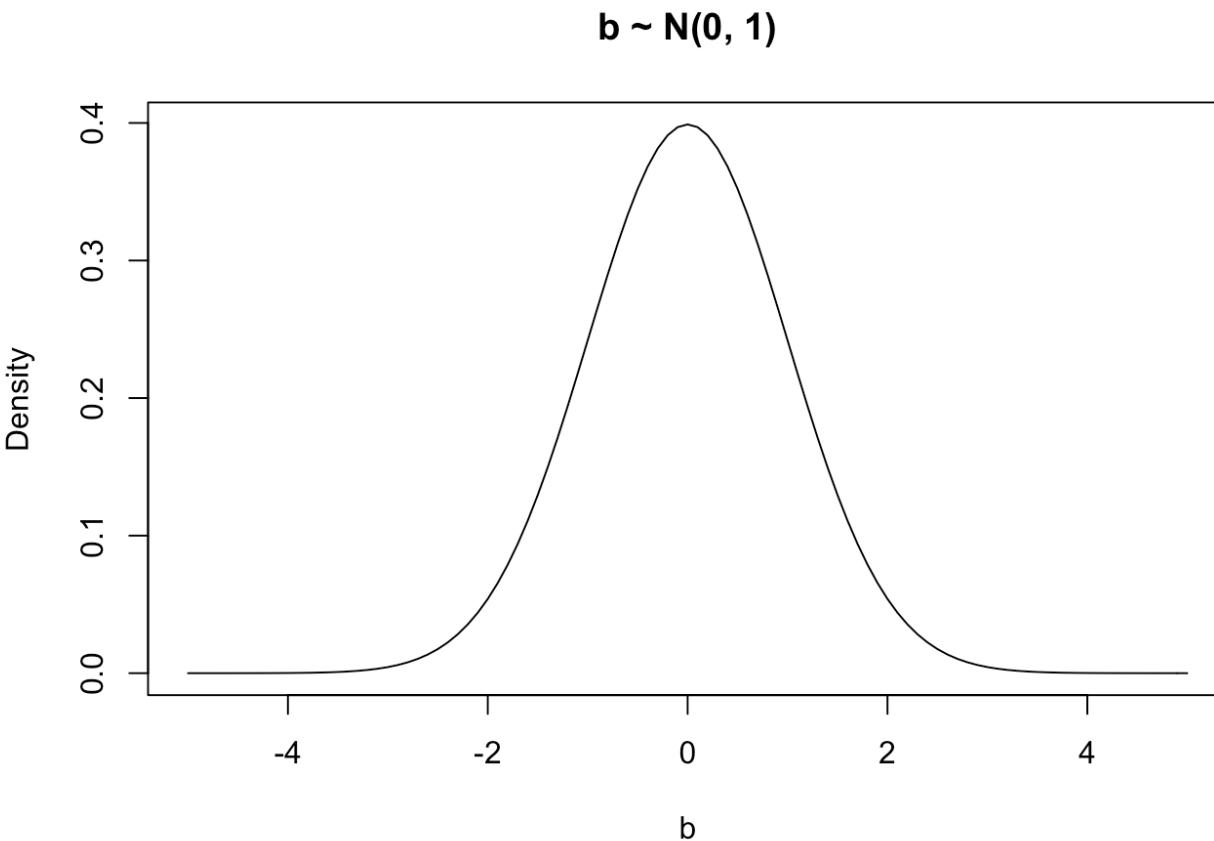
- Fewer/more flexible assumptions
- Use *prior* distributions to give your model an idea of what plausible parameter values would be
- Monte Carlo Markov Chain (MCMC) estimation (think lots of counting)
- Results are always in the form of distributions (can get point estimates via aggregation using various measures of central tendency)
- Credible intervals, Bayes Factors, probability of direction
- brms

Terms

- **Prior distribution:** represents our initial beliefs about the parameter estimate
 - Often conveyed in the form of a distribution centered around a value with some degree of spread
 - Example: a normal distribution centered around 0, with a standard deviation of 1; $\sim N(0, 1)$

Priors

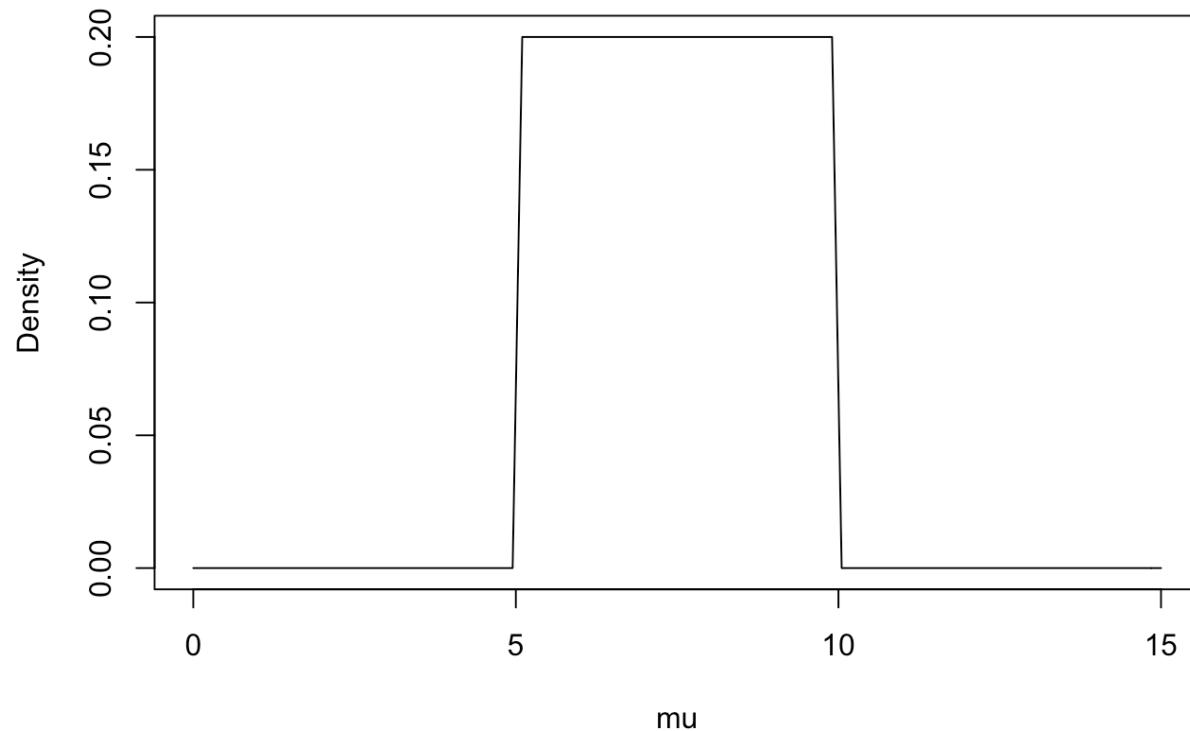
```
# normal  
curve(dnorm(x, 0, 1),from= -5, to = 5, main = "b ~ N(0, 1)", ylab = "Density", xlab = "b")
```



Priors

```
# uniform  
curve(dunif(mu, 5, 10), xname = "mu", from = 0, to = 15, main = "mu ~ uniform(5,10)", ylab = "Density")
```

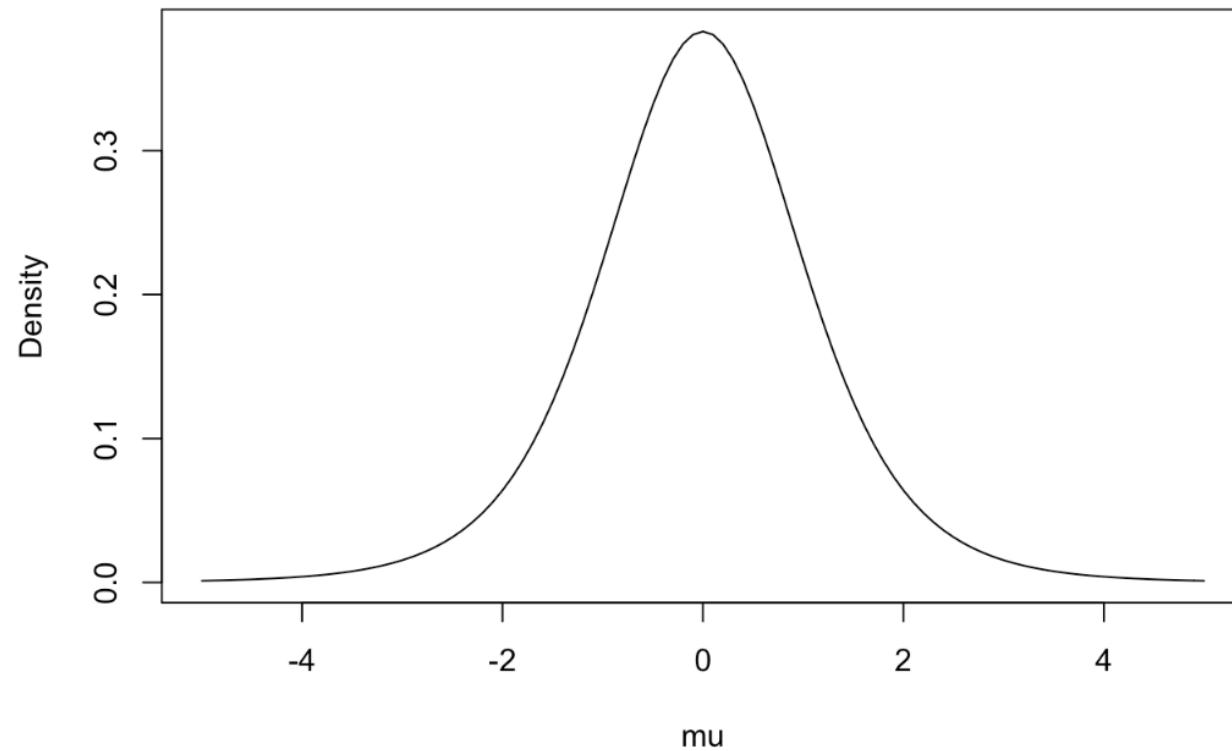
mu ~ uniform(5,10)



Priors

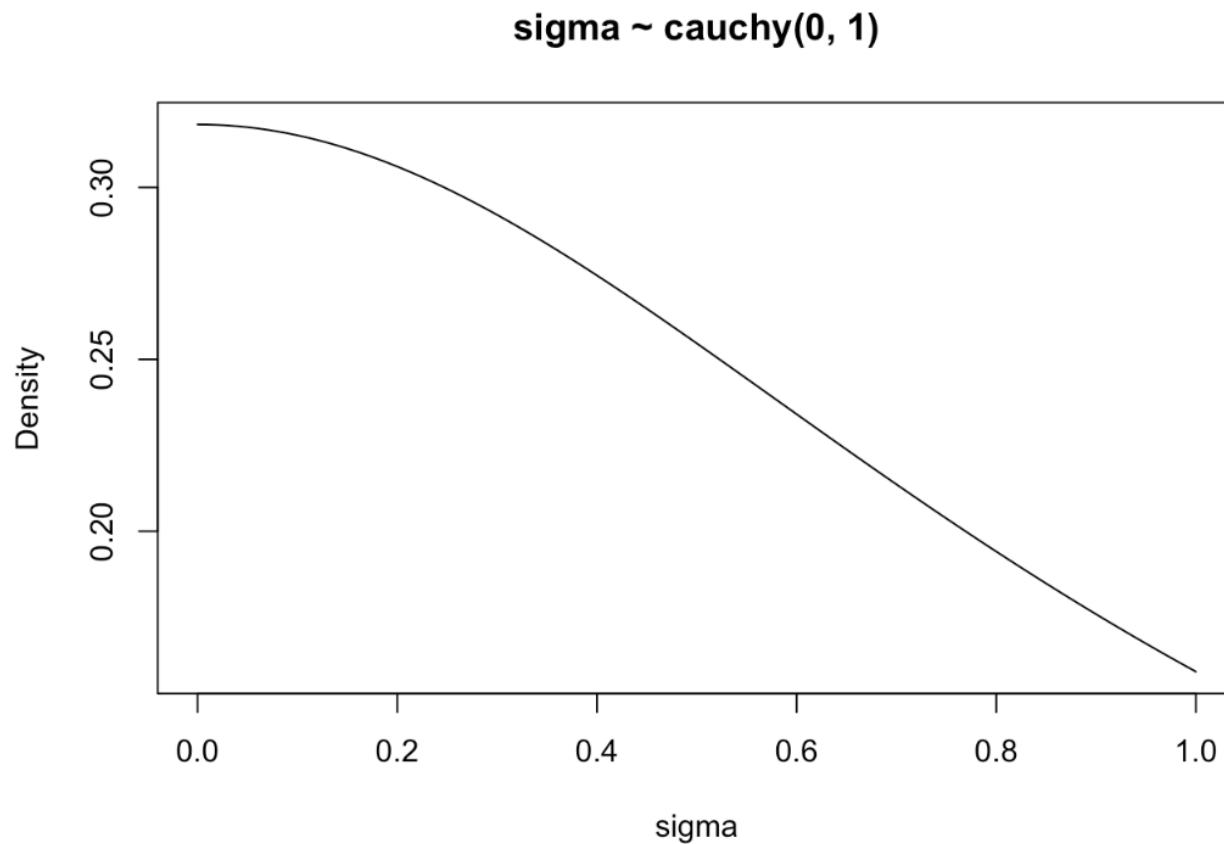
```
# t  
curve(dt(mu, 6, 0, .5), xname = "mu", from= -5, to = 5, main = "mu ~ t(6, 0, .5)", ylab = "Density")
```

$\text{mu} \sim t(6, 0, .5)$



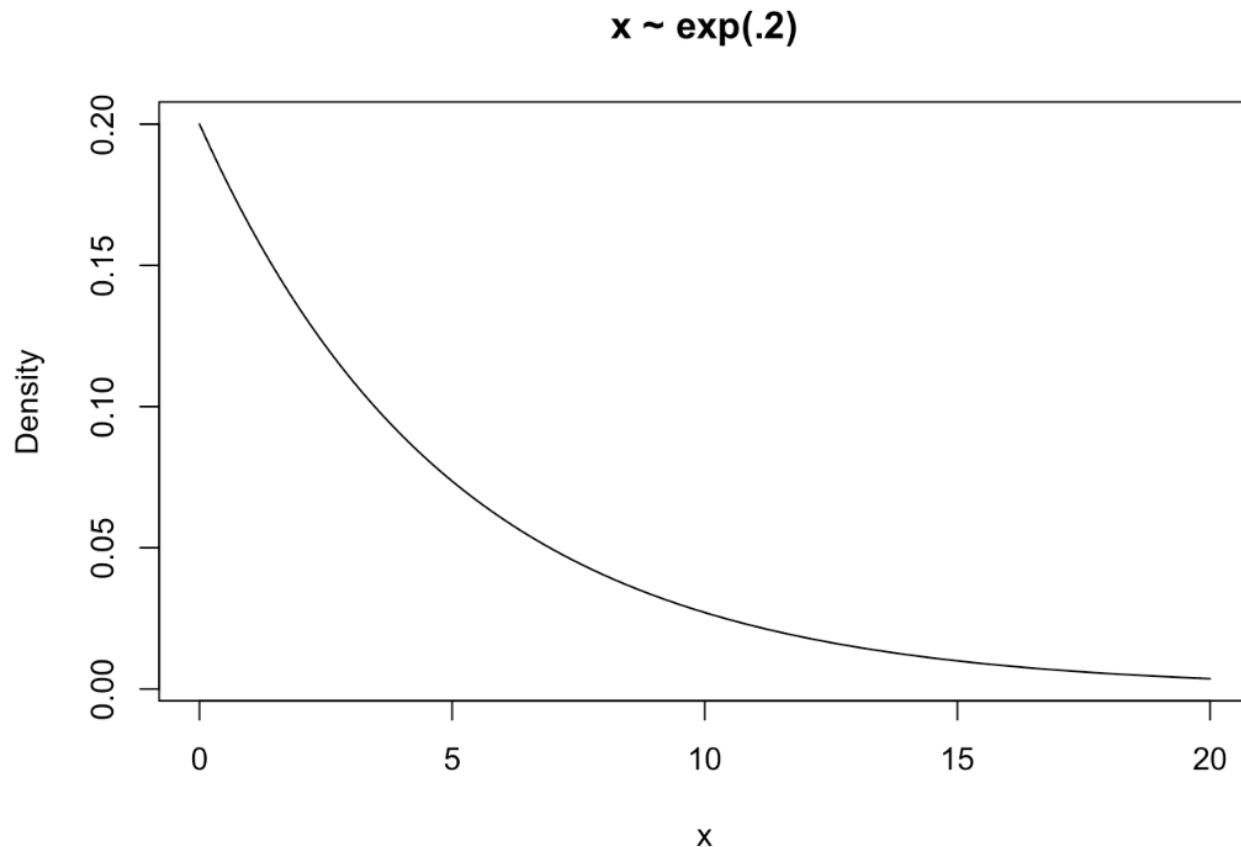
Priors

```
# cauchy
curve(dcauchy(sigma, 0, 1), xname = "sigma", from= 0, to = 1, main = "sigma ~ cauchy(0, 1)", ylab = "Density")
```



Priors

```
# exponential  
curve(dexp(x, .2), xname = "x", from= 0, to = 20, main = "x ~ exp(.2)", ylab = "Density")
```



Terms

- **Prior distribution:** represents our initial beliefs about the parameter estimate
 - Often conveyed in the form of a distribution centered around a value with some degree of spread
 - Example: a normal distribution centered around 0, with a standard deviation of 1; $\sim N(0, 1)$
- **Likelihood:** the relative number of ways the data could occur given our parameter estimates

Likelihood

- Mathematical function used to quantify the probability of different parameters
- $p(\text{Data} \mid \theta)$
- Binomial: probability of getting 3 successes out of 10 trials, assuming a probability of .5

$$p(k|N, p) = \binom{N}{k} \cdot p^k q^{N-k}$$

```
dbinom(3, size = 10, prob = .5)
```

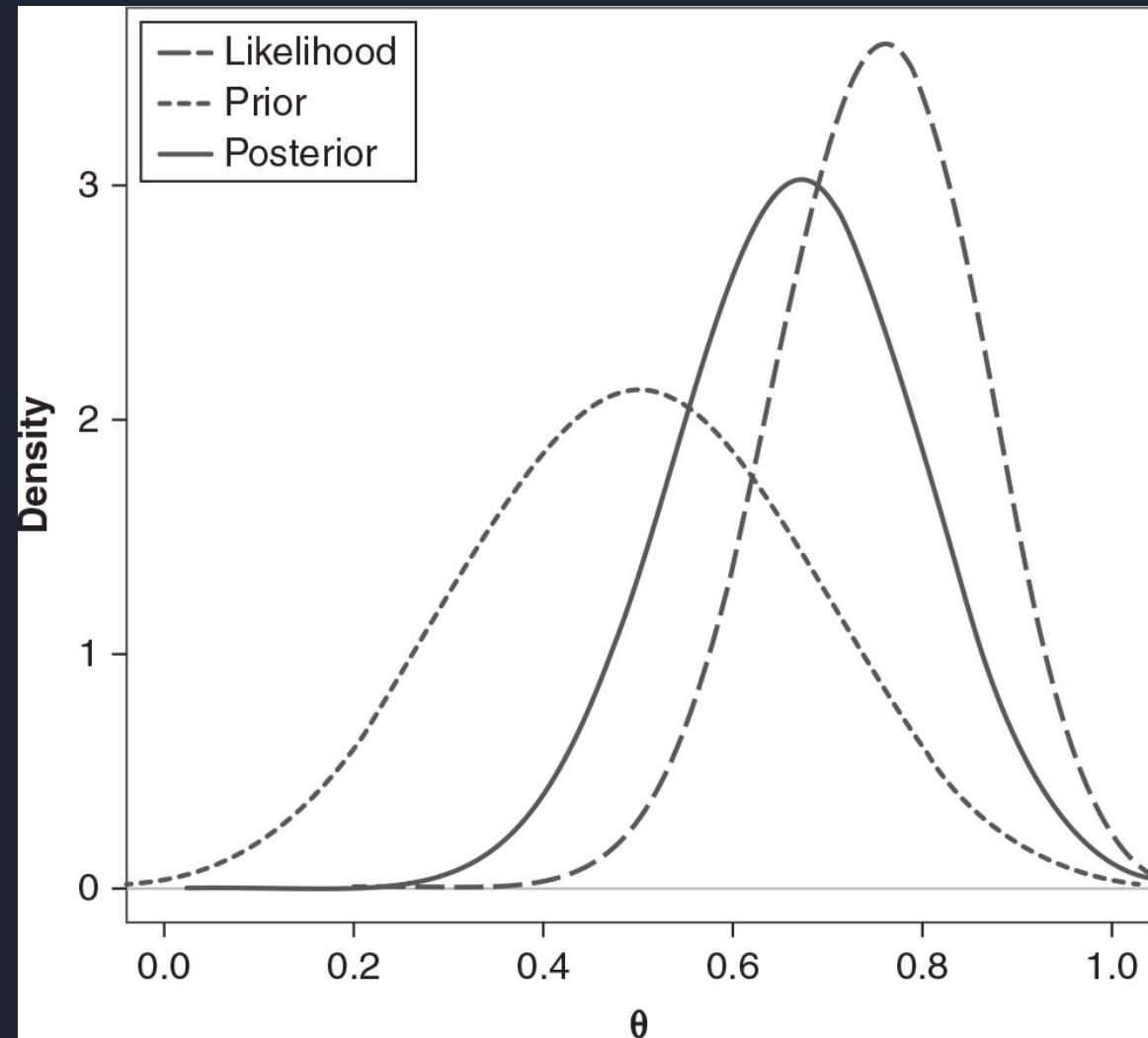
```
## [1] 0.1171875
```

Terms

- **Prior distribution:** represents our initial beliefs about the parameter estimate
 - Often conveyed in the form of a distribution centered around a value with some degree of spread
 - Example: a normal distribution centered around 0, with a standard deviation of 1; $\sim N(0, 1)$
- **Likelihood:** the relative number of ways the data could occur given our parameter estimates
- **Posterior distribution:** a weighted combination of the prior and the likelihood

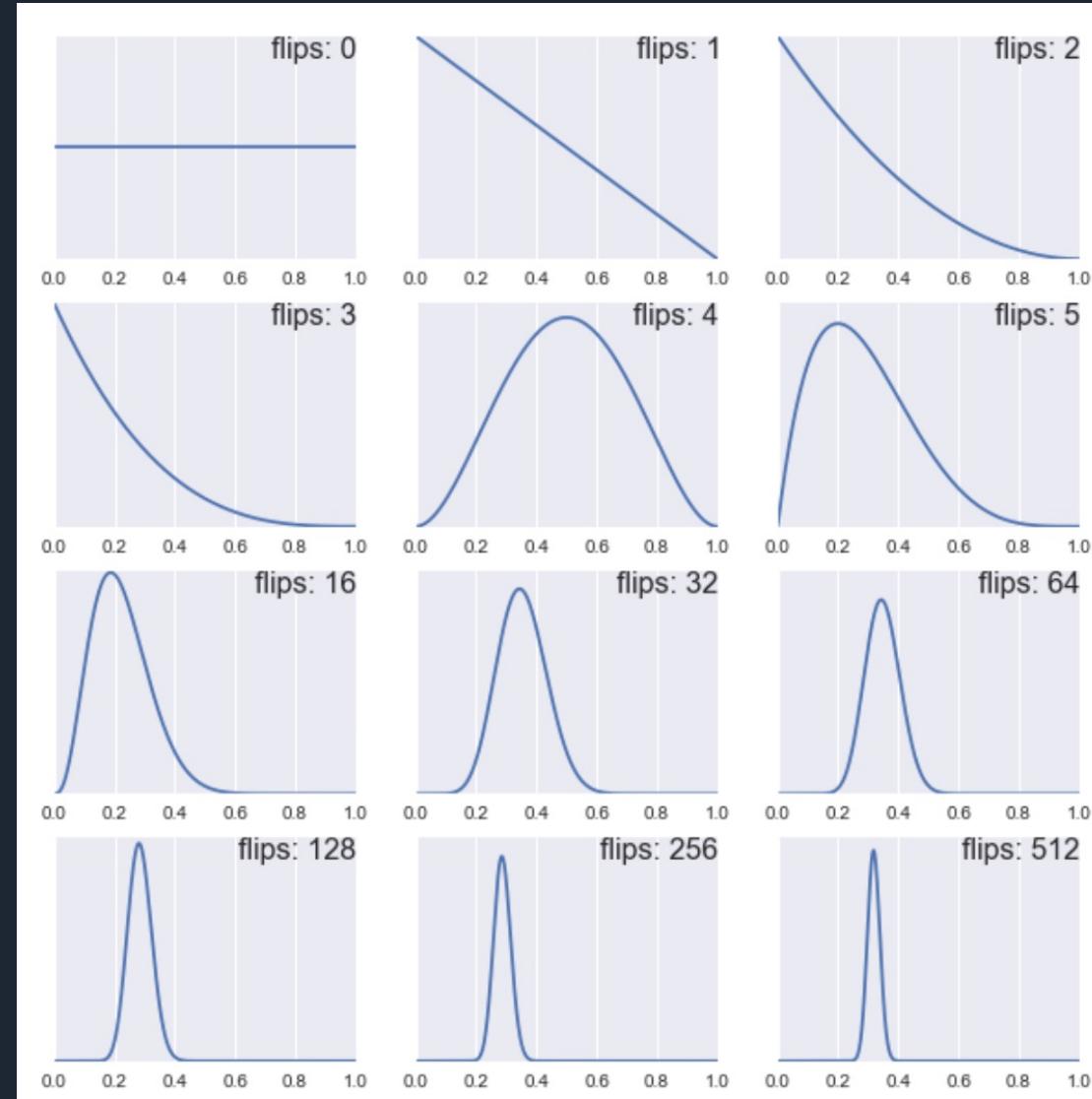
Posterior

- $p(\theta | \text{Data})$
- Updated prior distribution, after taking your data into account
- If you use flat priors, you can recreate estimates from frequentist estimation methods, as flat priors are implicit (such as in Maximum Likelihood)
 - The posterior is then equal to the likelihood
 - The likelihood is not appreciably different if you do ML or MCMC



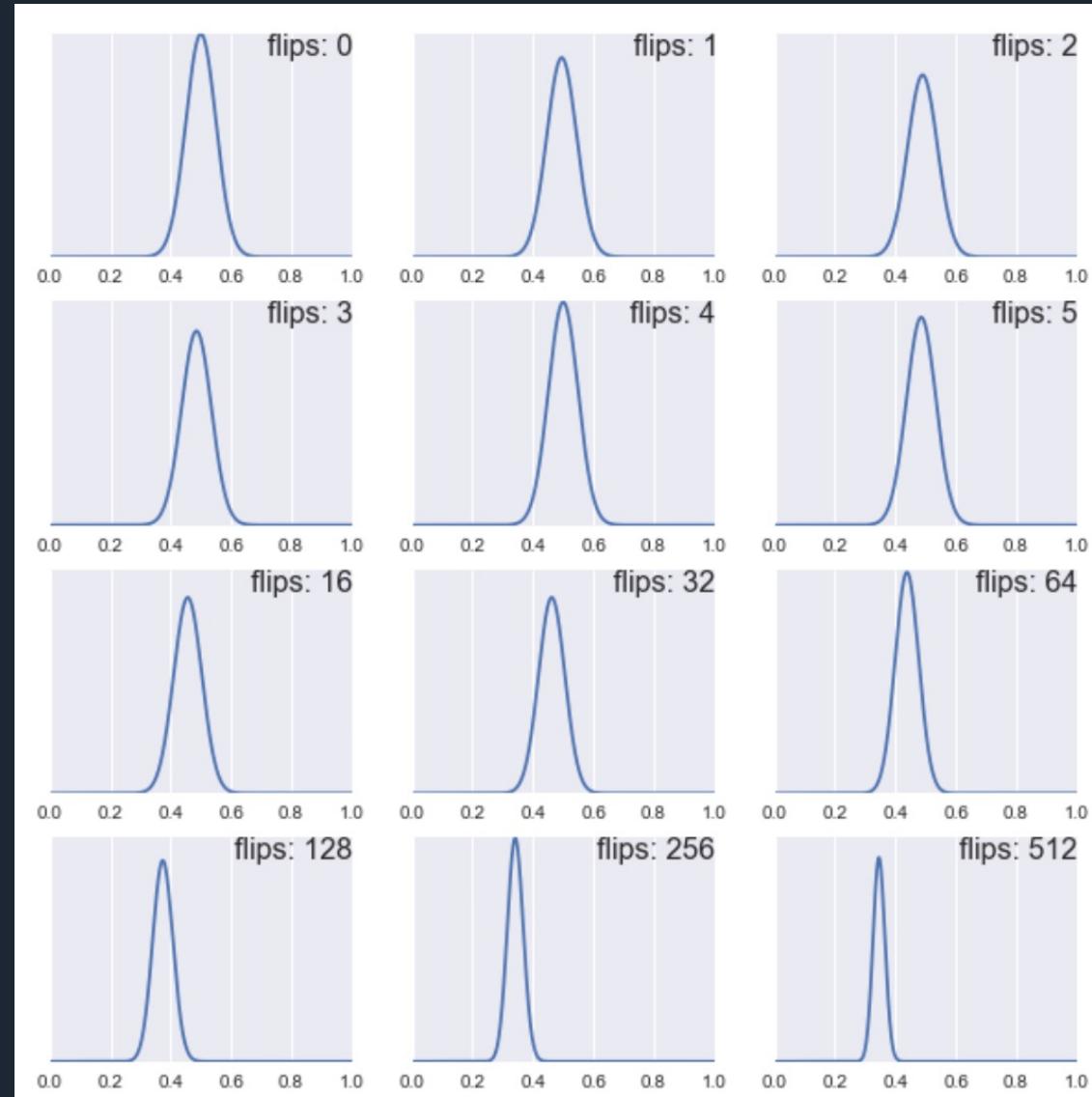
Prior → Posterior

Uniform prior



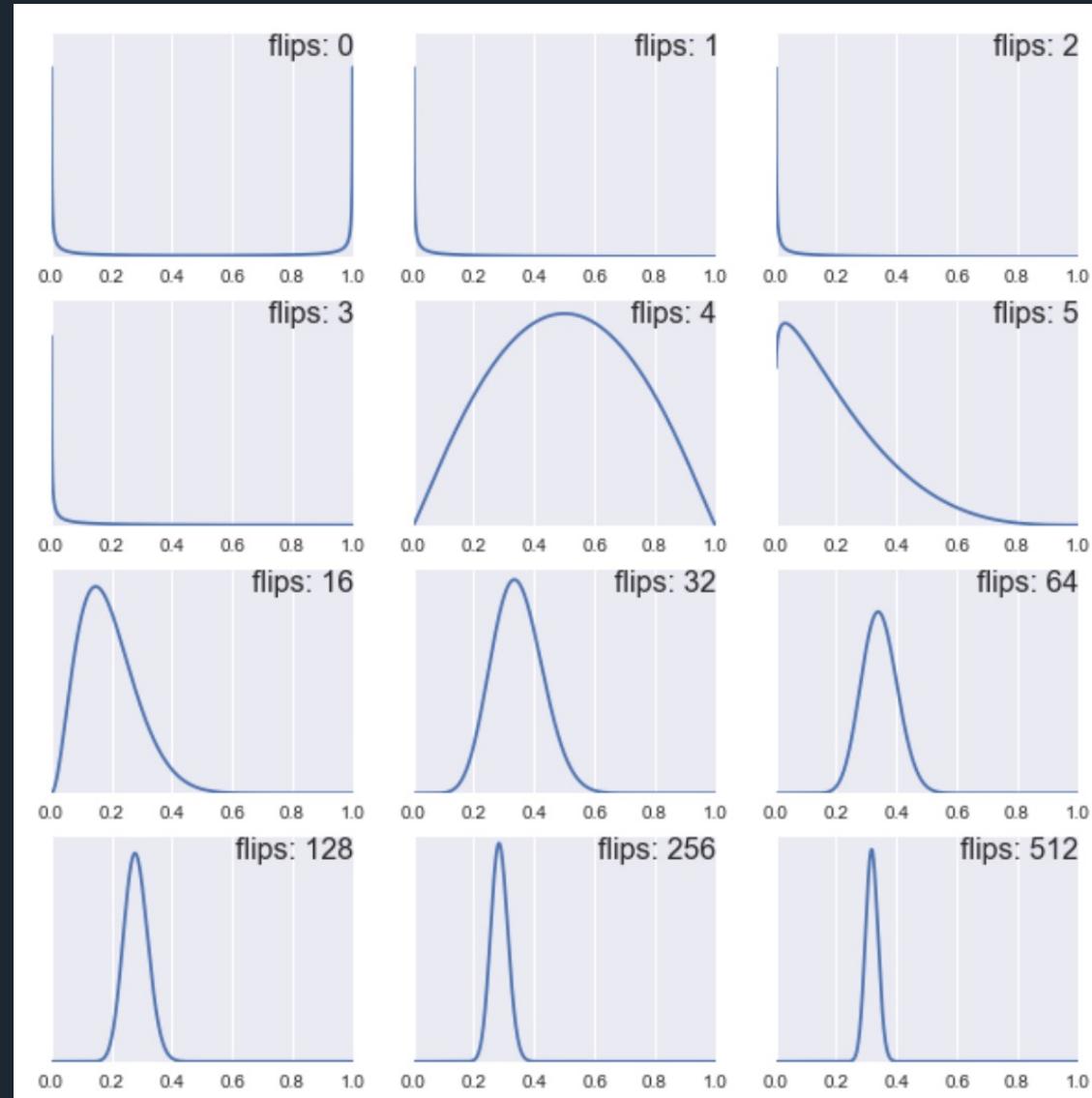
Prior → Posterior

Normal prior



Prior → Posterior

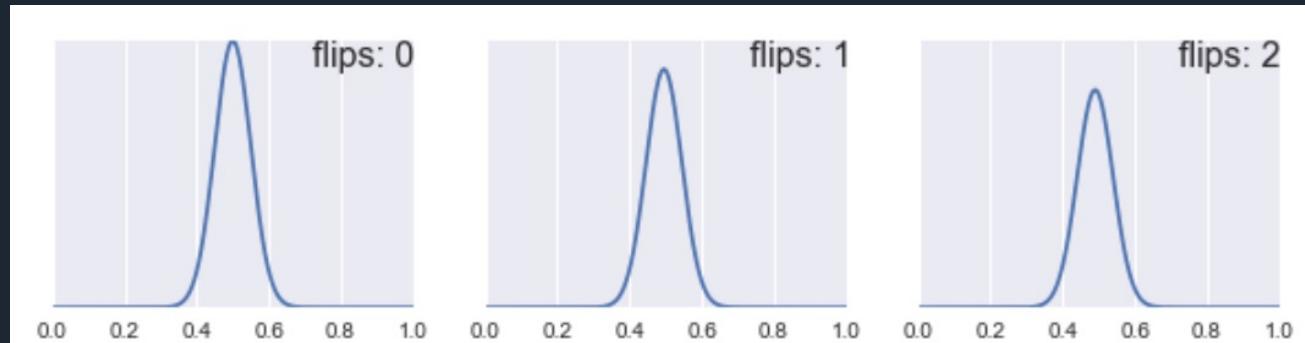
Beta prior



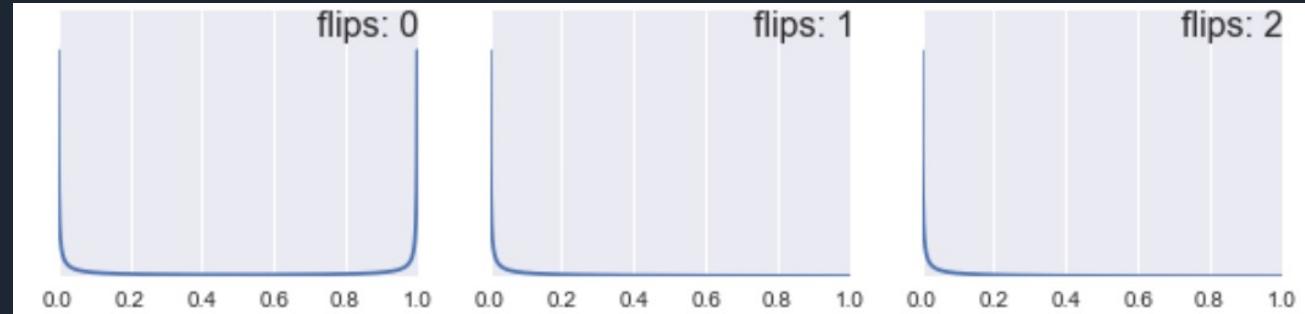
Uniform prior



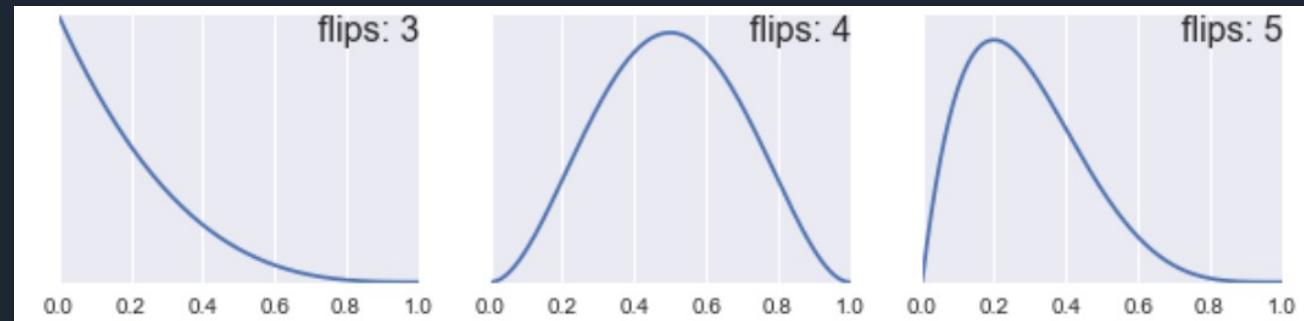
Normal prior



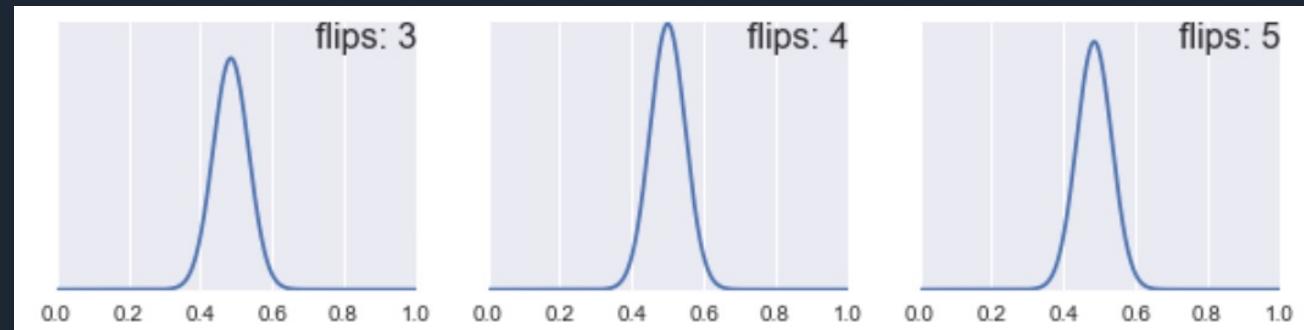
Beta prior



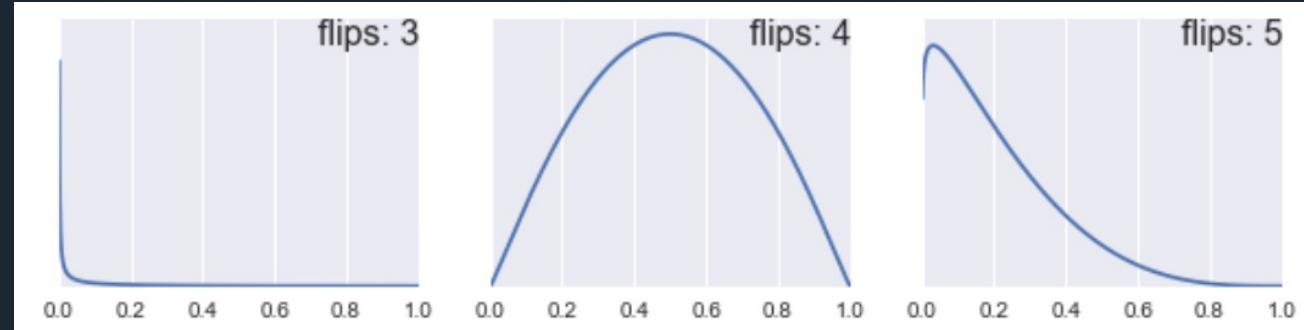
Uniform prior



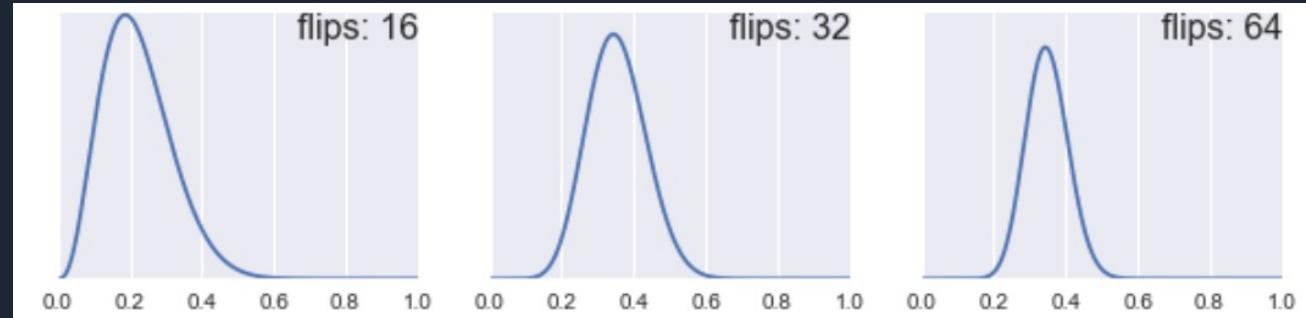
Normal prior



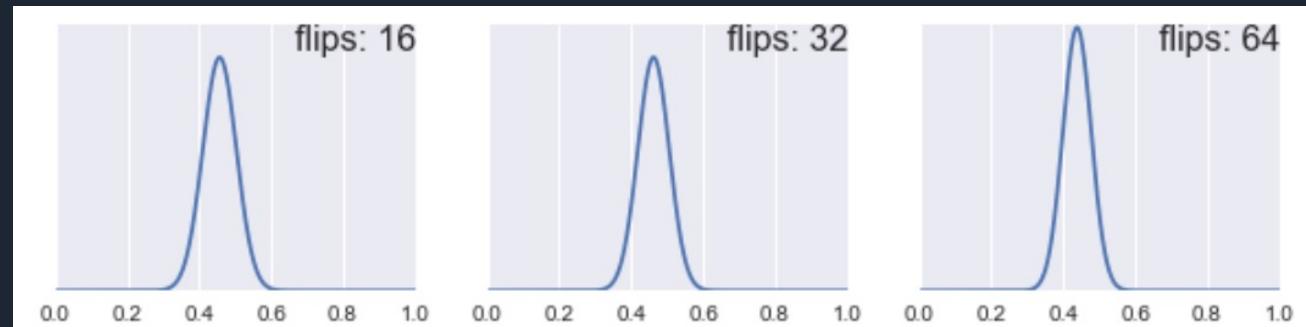
Beta prior



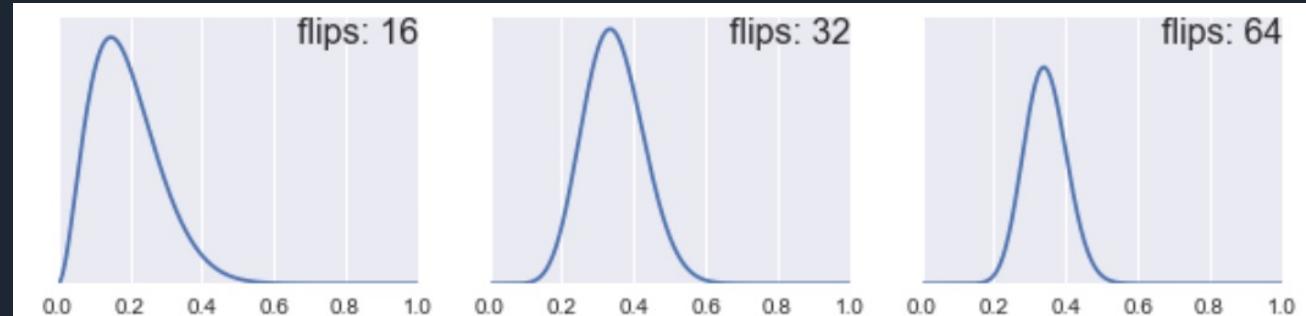
Uniform prior



Normal prior



Beta prior



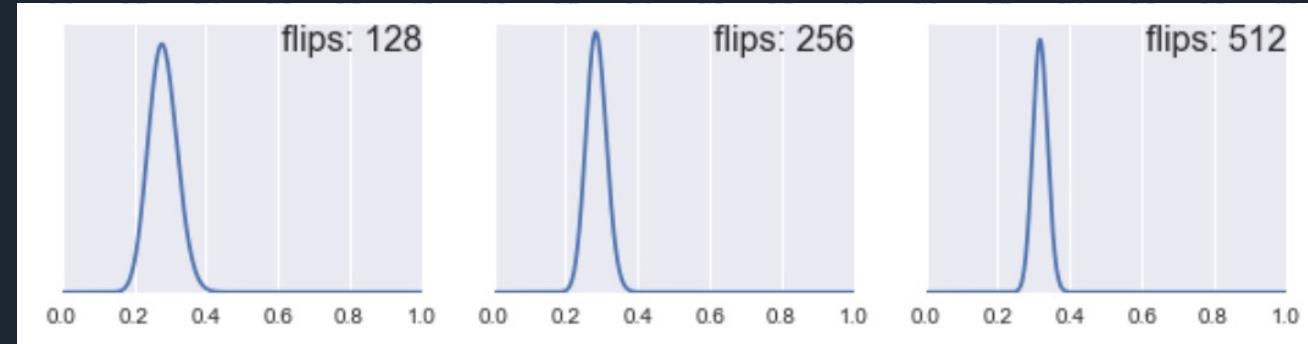
Uniform prior



Normal prior



Beta prior



Terms

- **Prior distribution:** represents our initial beliefs about the parameter estimate
 - Often conveyed in the form of a distribution centered around a value with some degree of spread
 - Example: a normal distribution centered around 0, with a standard deviation of 1; $\sim \mathcal{N}(0, 1)$
- **Likelihood:** the relative number of ways the data could occur given our parameter estimates
- **Posterior distribution:** a weighted combination of the prior and the likelihood
- **Bayes Theorem**
 - posterior \propto likelihood x prior
 - $p(\theta | \text{Data}) \propto p(\text{Data} | \theta) \times p(\theta)$

Bayes Theorem

$$p(\theta|data) \propto \frac{p(data|\theta) \times p(\theta)}{p(data)}$$

- $P(\theta|data)$ = posterior probability. Describes how confident we are that θ is true, given the observed data.
- $P(\theta)$ = prior probability. Describes how sure we were that θ was true, before we observed the data.
- $P(data|\theta)$ = likelihood. If you were to assume that θ is true, this is the probability that you would have observed your data.
- $P(data)$ is the average or marginal likelihood, sometimes called "the evidence"
 - The main purpose of this is to standardize the posterior so it integrates to 1

Terms

- **Prior distribution:** represents our initial beliefs about the parameter estimate
 - Often conveyed in the form of a distribution centered around a value with some degree of spread
 - Example: a normal distribution centered around 0, with a standard deviation of 1; $\sim N(0, 1)$
- **Likelihood:** the relative number of ways the data could occur given our parameter estimates
- **Posterior distribution:** a weighted combination of the prior and the likelihood
- **Bayes Theorem**
 - posterior \propto likelihood x prior
 - $p(\theta | \text{Data}) \propto p(\text{Data} | \theta) \times p(\theta)$
- **Credible intervals:** the probability the parameter falls within the bounds of your specified interval

A Quick Note About Priors...

- These are often the go-to (albeit uninformed) criticism about Bayesian analyses
- All models have researcher degrees of freedom
 - “Which variables should I include?”, “Should I model this interaction?”, “What form should my model be?”, “Should I *really* winsorize my data?” (probably not)
- Your priors are often “overwhelmed by the data” – ultimately have little impact on your model’s estimates besides giving it plausible values to look near
- You explicitly report your priors for models – they aren’t a secret
- Often chosen in a manner where they are weakly informative or regularize the estimates – avoids overfitting

Data

- Health & Retirement Study (HRS)
- $N = 2,707$
- Adults ($M_{age} = 63.5$)
- Single wave of data (i.e., cross-sectional)

What should we look at?

- Dependent variable: life satisfaction; range from 1-5 (LS_i)
- Independent variables:
 - Neuroticism scores at each wave; range from 0-3 (N_i)
 - Whether the participant is widowed at each wave; 0 = not widowed, 1 = widowed (W_i)
 - Gender; 0 = male, 1 = female ($Gender_i$)

Various Models

$$LS_i = b_0 + e_i$$

$$LS_i = b_0 + b_1 N_i + e_i$$

$$LS_i = b_0 + b_1 N_i + b_2 W_i + e_i$$

$$LS_i = b_0 + b_1 N_i + b_2 W_i + b_3 Gender_i + e_i$$

$$LS_i = b_0 + b_1 N_i + b_2 W_i + b_3 Gender_i + b_4 WxGender_i + e_i$$

via lm(): null model

```
freq_model_null <- lm(LS ~ 1, data = bayes_data)
summary(freq_model_null)

##
## Call:
## lm(formula = LS ~ 1, data = bayes_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8209 -0.4209  0.2458  0.6458  1.1791
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.82090   0.01676   228    <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8717 on 2705 degrees of freedom
## (1 observation deleted due to missingness)
```

$$LS_i = b_0 + e_i$$

via lm(): Neuroticism

```
freq_model_N <- lm(LS ~ 1 + N, data = bayes_data)
summary(freq_model_N)

##
## Call:
## lm(formula = LS ~ 1 + N, data = bayes_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3.2896 -0.4445  0.1372  0.6022  1.8724 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 4.28958   0.03168 135.42   <2e-16 ***
## N          -0.45347   0.02650 -17.11   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8282 on 2704 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.09774,    Adjusted R-squared:  0.09741 
## F-statistic: 292.9 on 1 and 2704 DF,  p-value: < 2.2e-16
```

$$LS_i = b_0 + b_1 N_i + e_i$$

via lm(): Neuroticism + Widowed

```
freq_model_widowed <- lm(LS ~ 1 + N + widowed, data = bayes_data)
summary(freq_model_widowed)

##
## Call:
## lm(formula = LS ~ 1 + N + widowed, data = bayes_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3.3044 -0.4519  0.1420  0.6137  1.8551 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 4.30445   0.03207 134.208 < 2e-16 ***
## N          -0.45250   0.02646 -17.099 < 2e-16 ***
## widowedWidowed -0.14214   0.05050  -2.815  0.00492 ** 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8271 on 2703 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1004, Adjusted R-squared:  0.09971 
## F-statistic: 150.8 on 2 and 2703 DF,  p-value: < 2.2e-16
```

$$LS_i = b_0 + b_1 N_i + b_2 W_i + e_i$$

via lm(): Neuroticism + Widowed + Gender

```
freq_model_gender <- lm(LS ~ 1 + N + widowed + gender, data = bayes_data)
summary(freq_model_gender)

##
## Call:
## lm(formula = LS ~ 1 + N + widowed + gender, data = bayes_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.3425 -0.4567  0.1540  0.6099  1.9070
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 4.26782   0.03597 118.650 < 2e-16 ***
## N          -0.45847   0.02658 -17.250 < 2e-16 ***
## widowedWidowed -0.16425   0.05142  -3.195  0.00142 ** 
## genderFemale  0.07469   0.03330   2.243  0.02499 *  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8265 on 2702 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.102,  Adjusted R-squared:  0.1011 
## F-statistic: 102.4 on 3 and 2702 DF,  p-value: < 2.2e-16
```

$$LS_i = b_0 + b_1 N_i + b_2 W_i + b_3 Gender_i + e_i$$

*via lm(): Neuroticism + Widowed*Gender*

```
freq_model_int <- lm(LS ~ 1 + N + widowed + gender + widowed:gender, data = bayes_data)
summary(freq_model_int)

##
## Call:
## lm(formula = LS ~ 1 + N + widowed + gender + widowed:gender,
##      data = bayes_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3.3414 -0.4575  0.1551  0.6076  1.9055 
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)              4.26937   0.03627 117.700  <2e-16 ***
## N                      -0.45850   0.02658 -17.248  <2e-16 ***
## widowedWidowed          -0.20593   0.13486 -1.527   0.1269    
## genderFemale              0.07203   0.03424  2.104   0.0355 *  
## widowedWidowed:genderFemale 0.04877   0.14588  0.334   0.7381    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8266 on 2701 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1021, Adjusted R-squared:  0.1008 
## F-statistic: 76.77 on 4 and 2701 DF,  p-value: < 2.2e-16
```

$$LS_i = b_0 + b_1 N_i + b_2 W_i + b_3 Gender_i + b_4 WxGender_i + e_i$$

Model Comparison

```
anova(freq_model_null, freq_model_N, freq_model_widowed, freq_model_gender, freq_model_int)

## Analysis of Variance Table
##
## Model 1: LS ~ 1
## Model 2: LS ~ 1 + N
## Model 3: LS ~ 1 + N + widowed
## Model 4: LS ~ 1 + N + widowed + gender
## Model 5: LS ~ 1 + N + widowed + gender + widowed:gender
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)
## 1    2705 2055.4
## 2    2704 1854.5  1    200.895 294.0126 < 2.2e-16 ***
## 3    2703 1849.1  1     5.419   7.9314  0.004894 **
## 4    2702 1845.6  1     3.436   5.0288  0.025010 *
## 5    2701 1845.6  1     0.076   0.1118  0.738140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

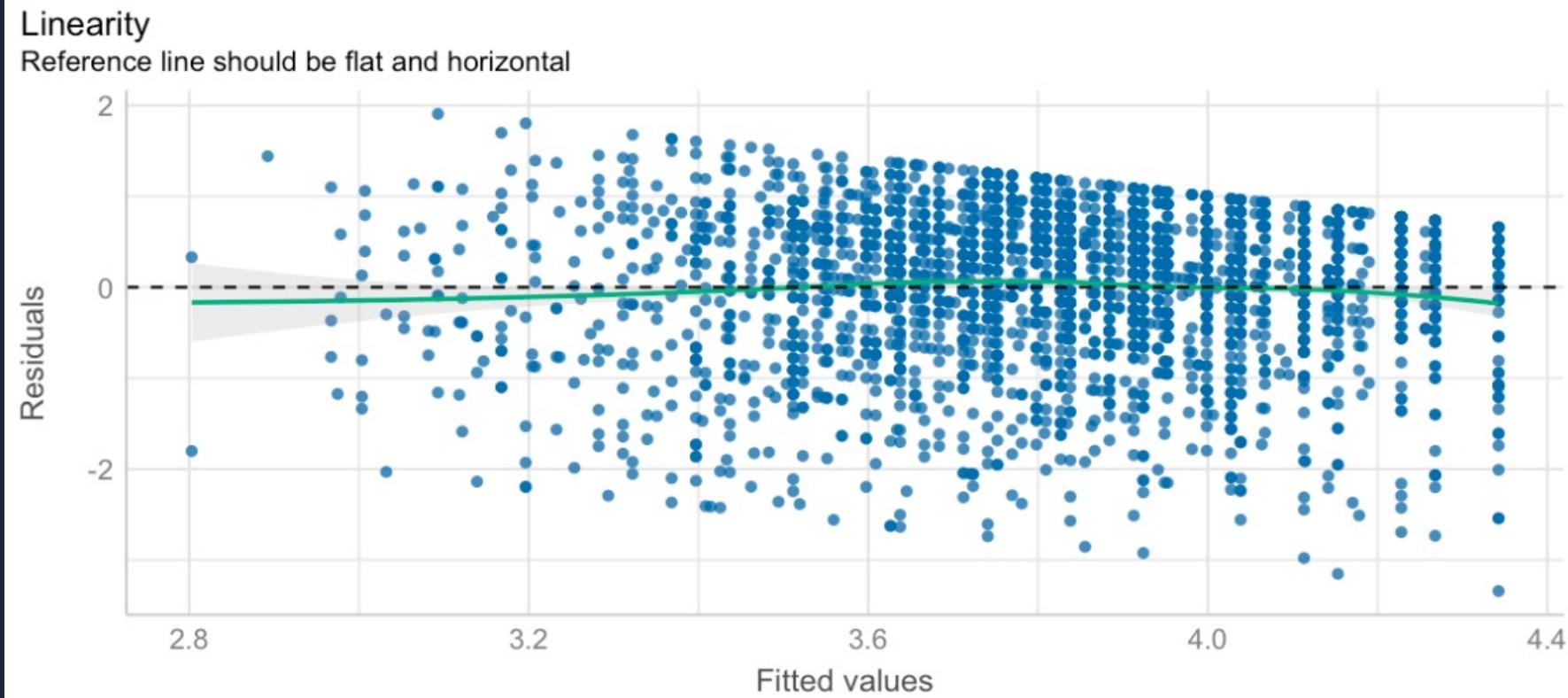
via lm(): Neuroticism + Widowed + Gender

```
freq_model_gender <- lm(LS ~ 1 + N + widowed + gender, data = bayes_data)
summary(freq_model_gender)

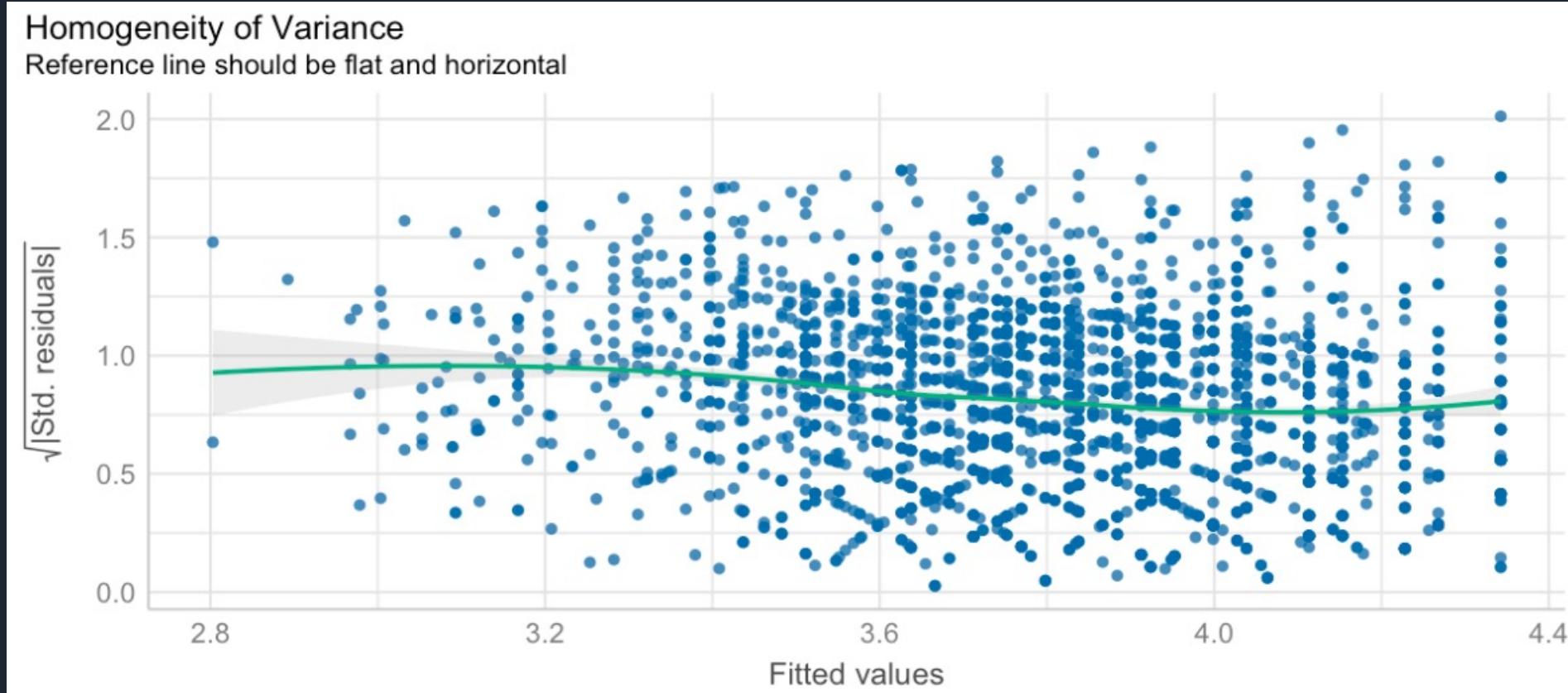
##
## Call:
## lm(formula = LS ~ 1 + N + widowed + gender, data = bayes_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.3425 -0.4567  0.1540  0.6099  1.9070
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 4.26782   0.03597 118.650 < 2e-16 ***
## N          -0.45847   0.02658 -17.250 < 2e-16 ***
## widowedWidowed -0.16425   0.05142  -3.195  0.00142 ** 
## genderFemale  0.07469   0.03330   2.243  0.02499 *  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8265 on 2702 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.102,  Adjusted R-squared:  0.1011 
## F-statistic: 102.4 on 3 and 2702 DF,  p-value: < 2.2e-16
```

$$LS_i = b_0 + b_1 N_i + b_2 W_i + b_3 Gender_i + e_i$$

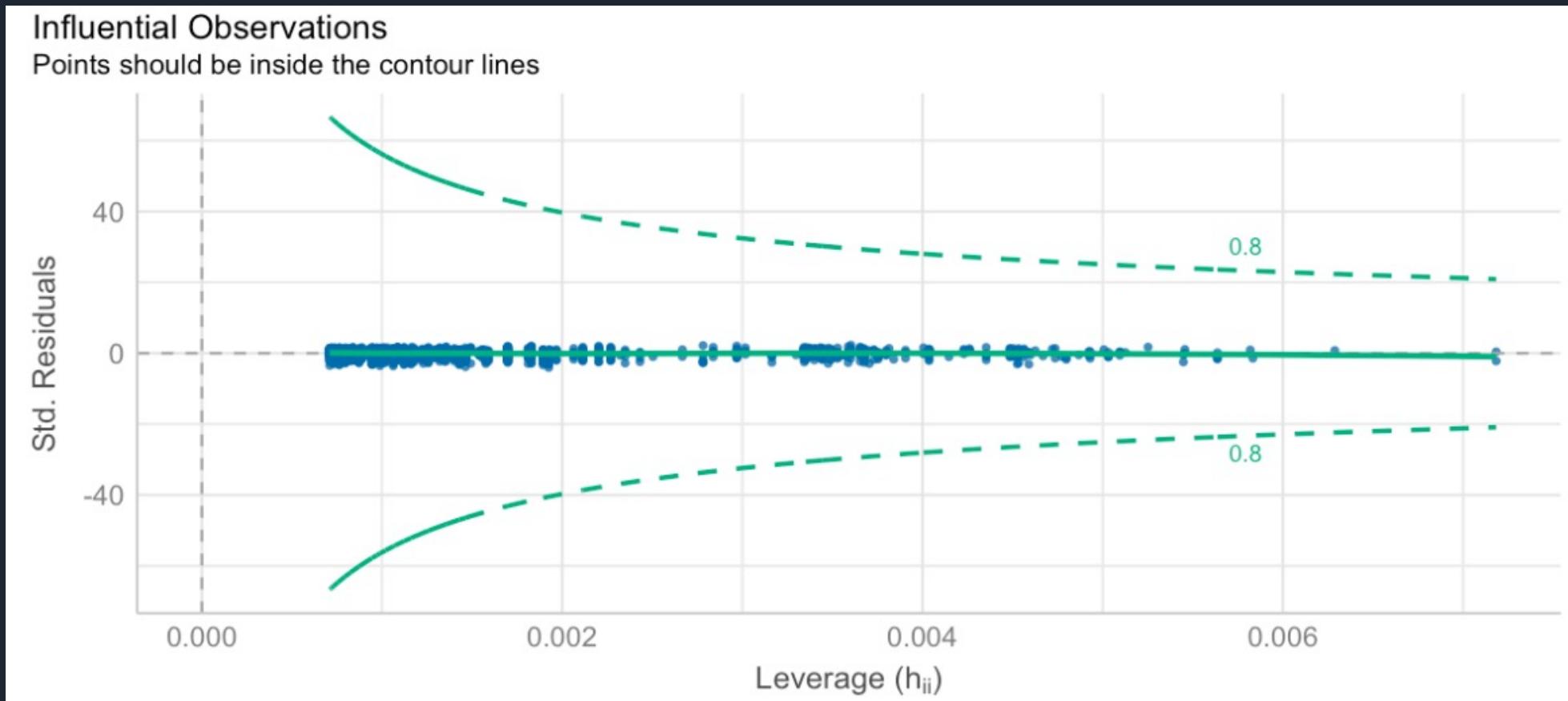
Check Assumptions



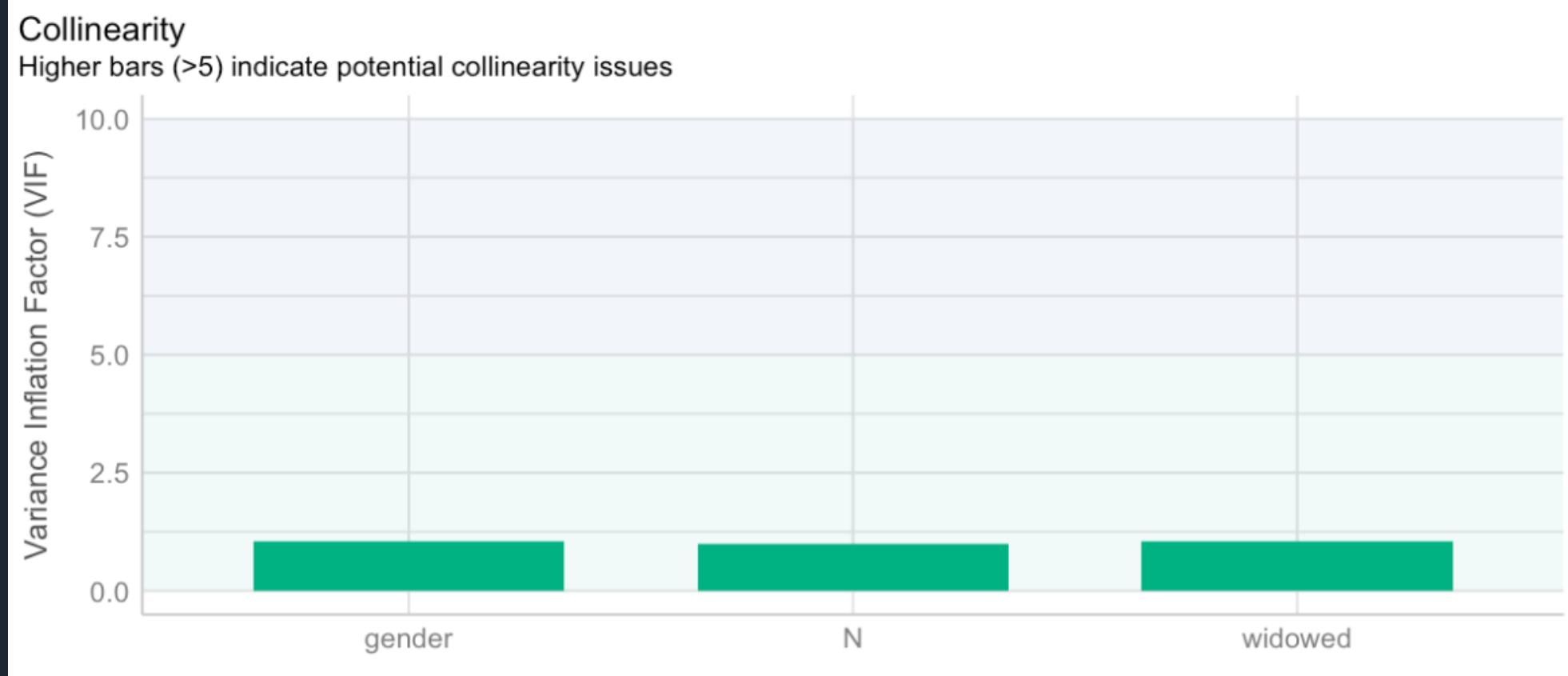
Check Assumptions



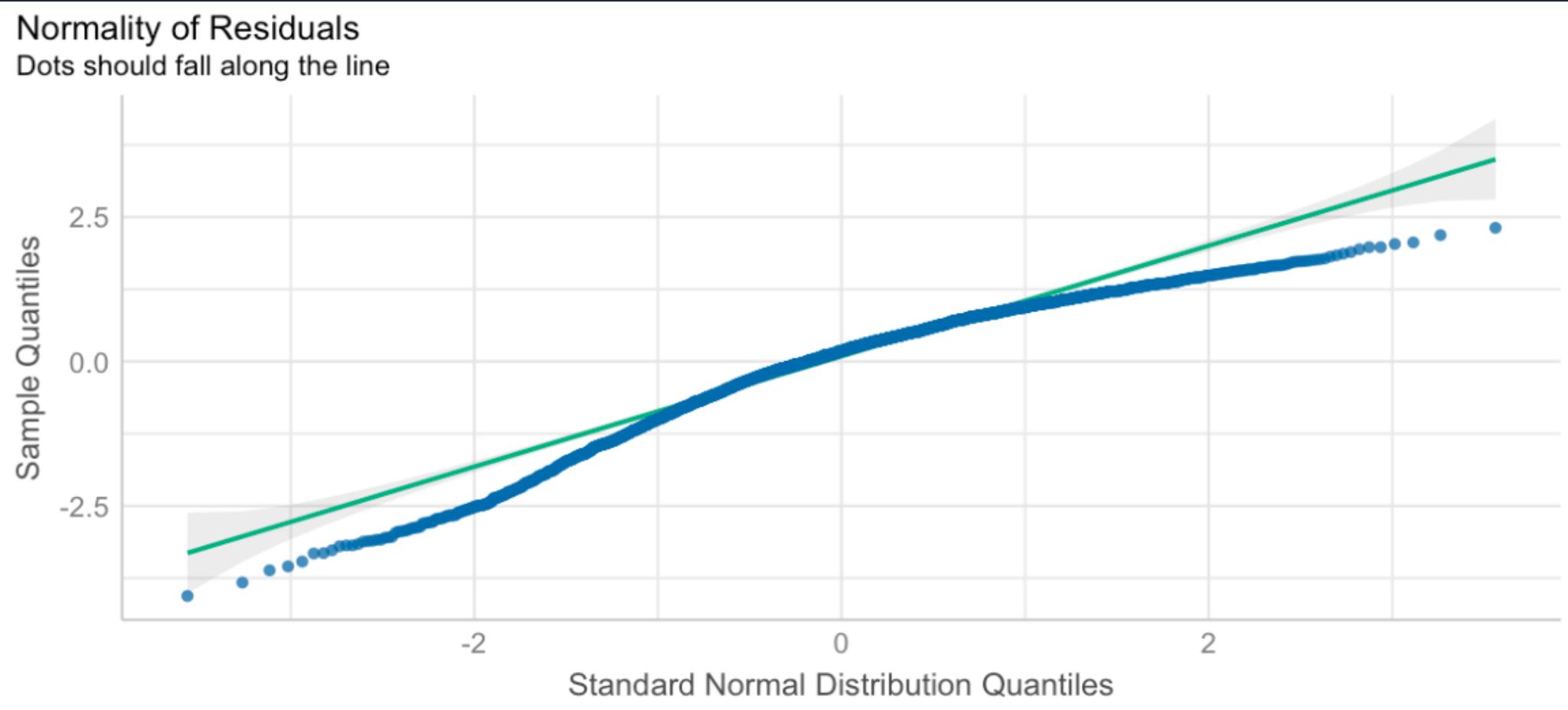
Check Assumptions



Check Assumptions



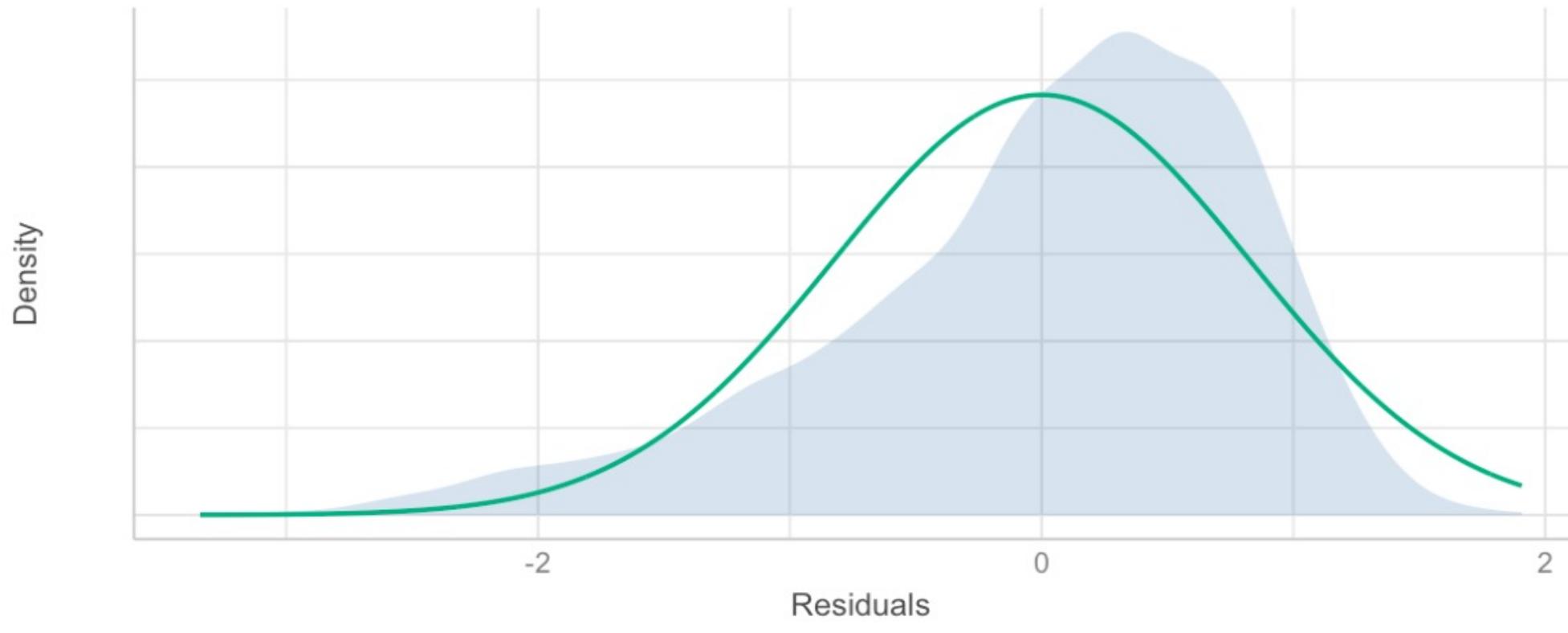
Check Assumptions



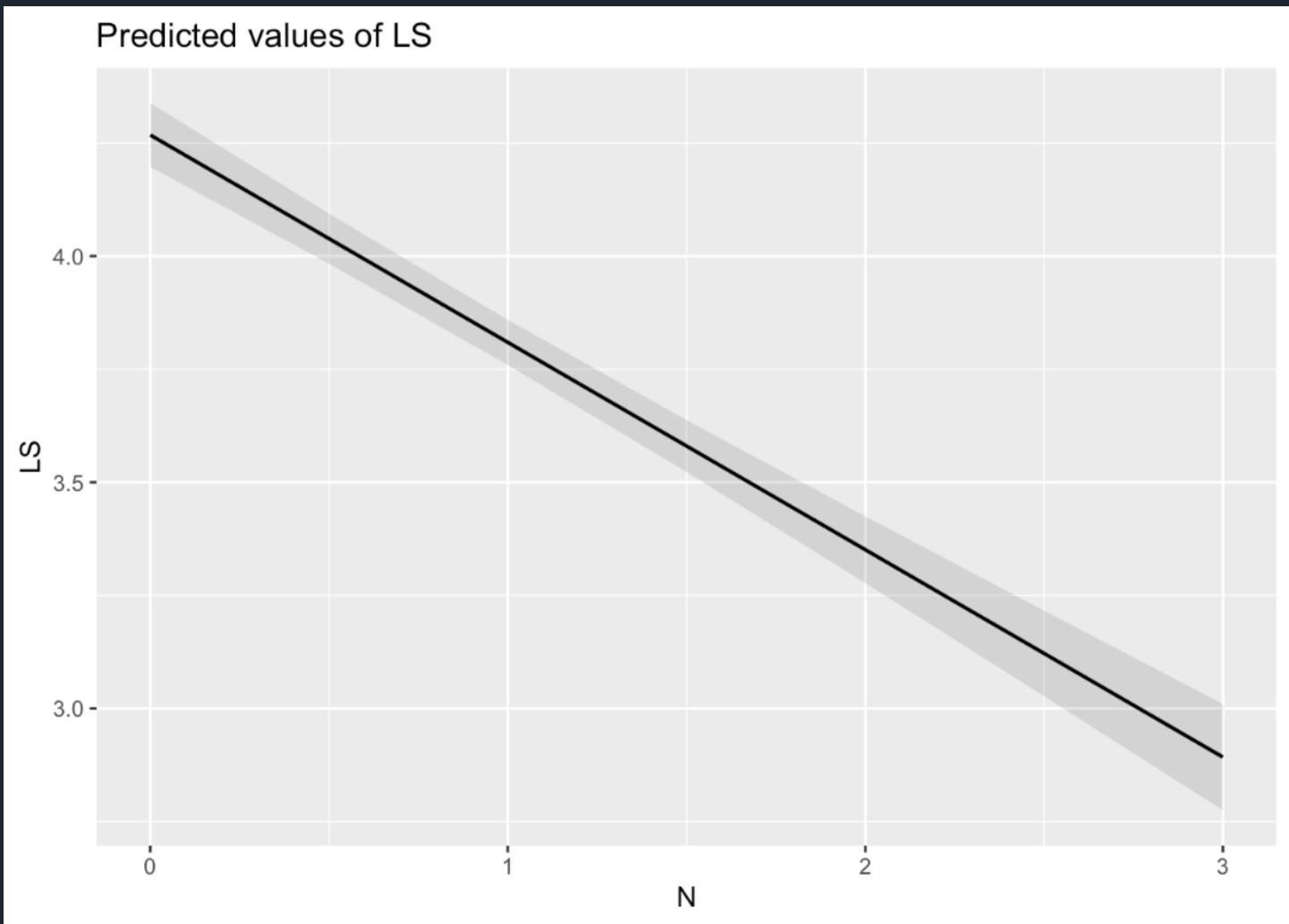
Check Assumptions

Normality of Residuals

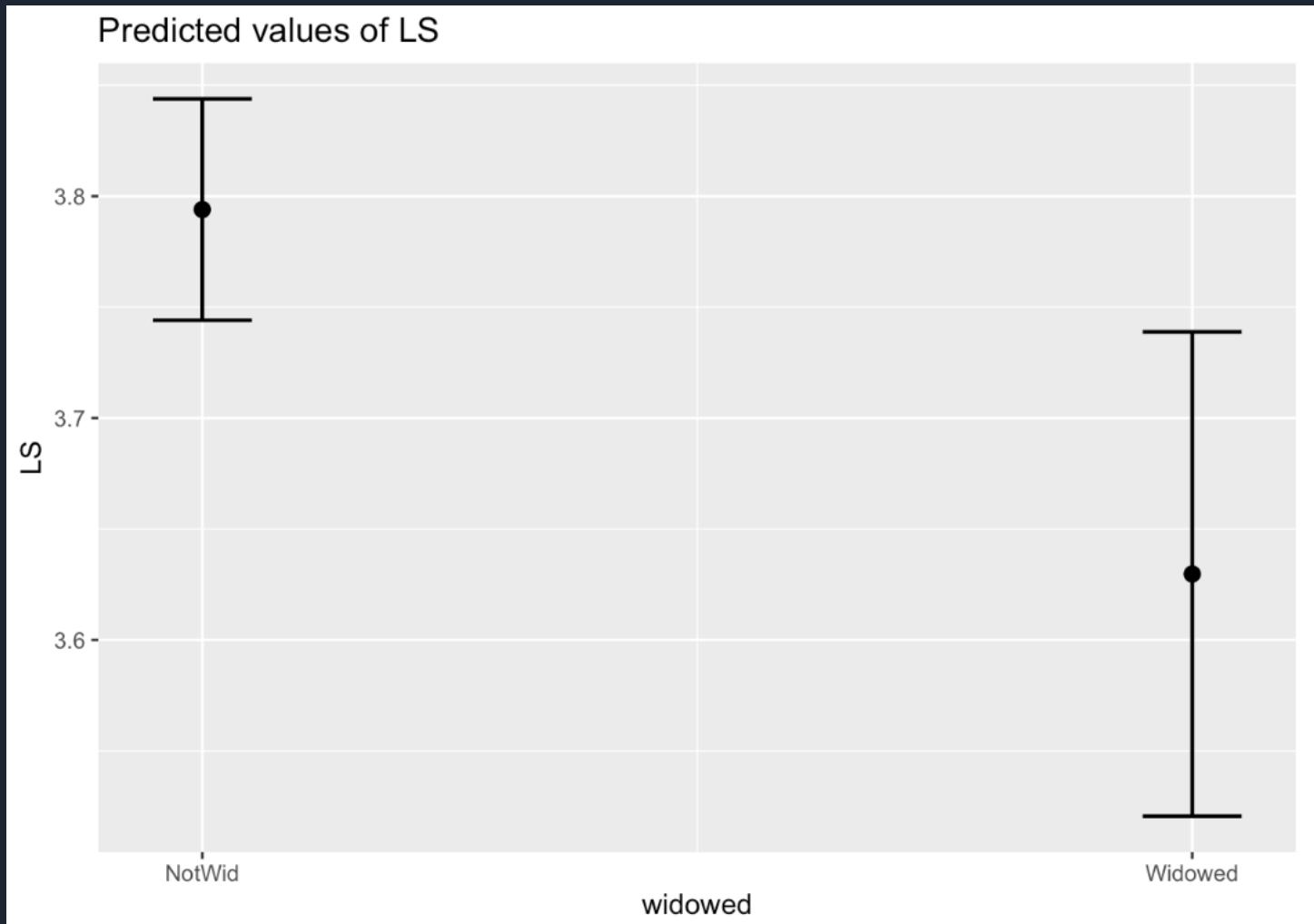
Distribution should be close to the normal curve



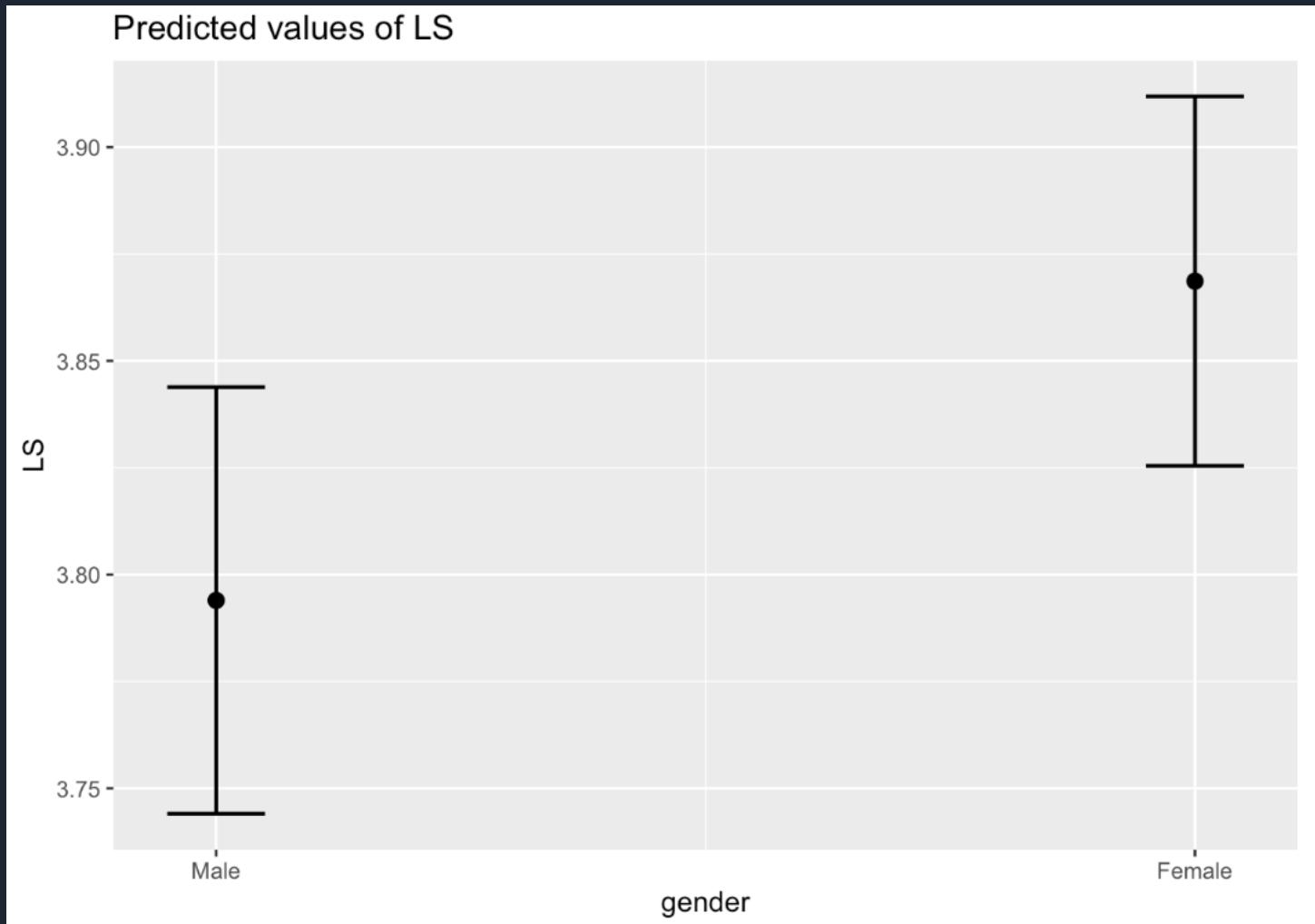
Plot of lm()



Plot of lm()



Plot of lm()



In brms

$$LS_i = b_0 + b_1 N_i + b_2 W_i + b_3 Gender_i + b_4 WxGender_i + e_i$$

- Same model, same specification
- Add in your priors for your parameter estimates
- Depending on the complexity of the model, size of dataset, and number of iterations you specify, it take some time to run the model

Priors for Our Model

- Can use the default priors provided by brms or specify your own
 - For defaults, use `get_prior()`
- Our intercept is the average life satisfaction score for a male participant with the minimum possible neuroticism score who is not widowed
 - Can calculate this value in your dataset yourself
 - Average: 4.22
- Can check your priors before running your model to see if they are appropriate and modify as needed

Default Priors

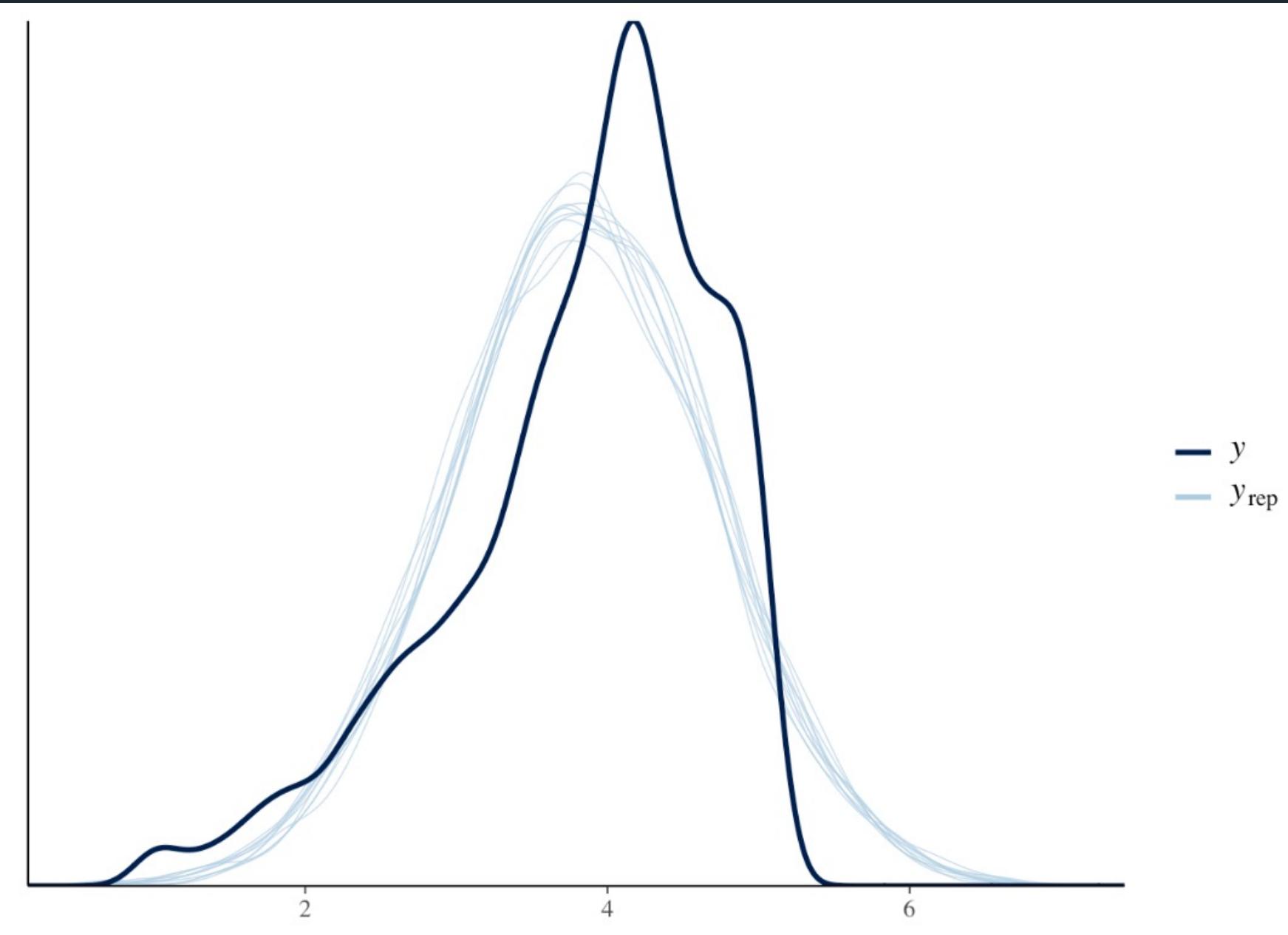
```
# default
prior <- get_prior(bf(LS ~ 1 + N + widowed + gender), data = bayes_data)
```

	prior	class	coef	group	resp	dpar	nLpar	bound	source
1		b							default
2		b	genderFemale						default
3		b	N						default
4		b	widowedWidowed						default
5	student_t(3, 4.1, 2.5)	Intercept							default
6	student_t(3, 0, 2.5)	sigma							default

Checking Priors

```
## ND  
brms_priors_ND <-  
  brm(family = gaussian,  
        LS ~ 1 + N + widowed + gender,  
        prior = c(prior(normal(4, .5), class = Intercept),  
                  prior(normal(0, .5), class = b),  
                  prior(exponential(1), class = sigma)),  
        chains = 4, cores = 4, iter = 4000,  
        data = bayes_data,  
        sample_prior = T)
```

Elapsed Time: 0.243391 seconds (Warm-up)
0.275857 seconds (Sampling)
0.519248 seconds (Total)



brms prior output

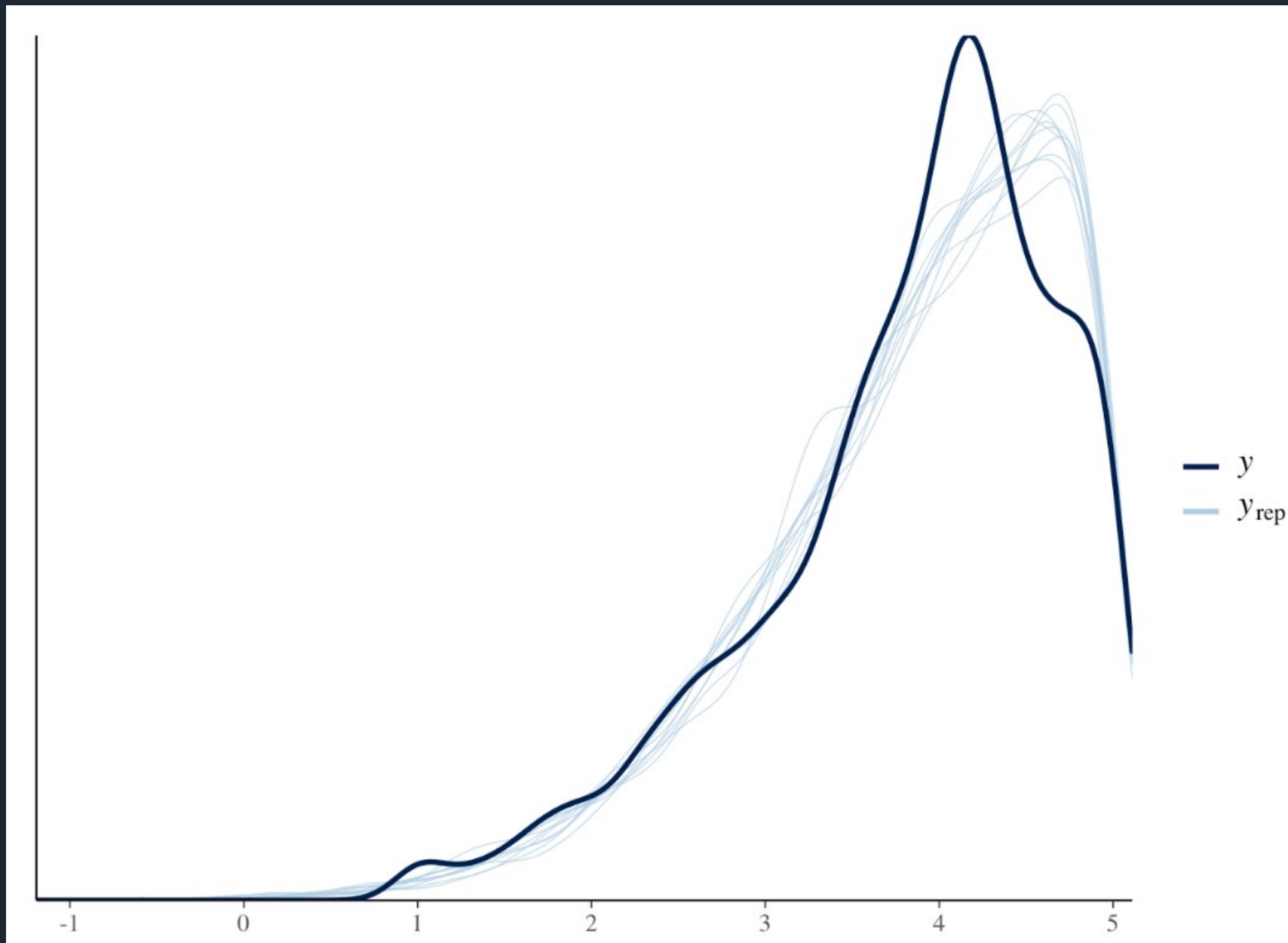
```
summary(brms_priors_ND)

## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LS ~ 1 + N + widowed + gender
## Data: bayes_data (Number of observations: 2706)
## Draws: 4 chains, each with iter = 4000; warmup = 2000; thin = 1;
##        total post-warmup draws = 8000
##
## Population-Level Effects:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      4.27     0.04    4.19    4.34 1.00   10175    5778
## N            -0.46     0.03   -0.51   -0.41 1.00    9652    5881
## widowedWidowed -0.16     0.05   -0.26   -0.06 1.00    9900    6650
## genderFemale    0.07     0.03    0.01    0.14 1.00    9576    6479
##
## Family Specific Parameters:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.83     0.01    0.80    0.85 1.00    9862    6165
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Checking Priors

```
## skew normal
brms_priors_skewND <-
  brm(family = skew_normal,
       LS ~ 1 + N + widowed + gender,
       prior = c(prior(normal(4, .5), class = Intercept),
                  prior(normal(0, .5), class = b),
                  prior(normal(-20, 10), class = alpha),
                  prior(exponential(1), class = sigma)),
       chains = 4, cores = 4, iter = 4000,
       data = bayes_data,
       sample_prior = T)
```

Elapsed Time: 26.2404 seconds (Warm-up)
26.573 seconds (Sampling)
52.8134 seconds (Total)



```
summary(brms_priors_skewND)

## Family: skew_normal
## Links: mu = identity; sigma = identity; alpha = identity
## Formula: LS ~ 1 + N + widowed + gender
## Data: bayes_data (Number of observations: 2706)
## Draws: 4 chains, each with iter = 4000; warmup = 2000; thin = 1;
##        total post-warmup draws = 8000
##
## Population-Level Effects:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      3.87     0.02    3.83    3.90 1.00    4140    4341
## N            -0.03     0.01   -0.04   -0.01 1.00    4697    4035
## widowedWidowed -0.02     0.01   -0.05    0.01 1.00    5574    4342
## genderFemale    0.01     0.01   -0.00    0.03 1.00    5358    4920
##
## Family Specific Parameters:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.89     0.01     0.87    0.91 1.00    3843    4177
## alpha     -63.15    7.87   -78.36   -47.36 1.00    4404    4523
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Specifying the Model

```
# only main effects
brms_model_ND <-
  brm(family = gaussian,
       LS ~ 1 + N + widowed + gender,
       prior = c(prior(normal(4, .5), class = Intercept),
                  prior(normal(0, .5), class = b),
                  prior(exponential(1), class = sigma)),
       chains = 4, cores = 4, iter = 6000, warmup = 3000,
       data = bayes_data,
       file = "brms_model_ND")
```

```
summary(brms_model_ND)

## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LS ~ 1 + N + widowed + gender
## Data: bayes_data (Number of observations: 2706)
## Draws: 4 chains, each with iter = 6000; warmup = 3000; thin = 1;
##        total post-warmup draws = 12000
##
## Population-Level Effects:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      4.27     0.04     4.20     4.34 1.00    14449    9034
## N            -0.46     0.03    -0.51    -0.41 1.00    14079    9626
## widowedWidowed -0.16     0.05    -0.26    -0.06 1.00    14129    9449
## genderFemale    0.07     0.03     0.01     0.14 1.00    13411    9716
##
## Family Specific Parameters:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.83     0.01     0.81     0.85 1.00    14480    9533
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

brms output

Elapsed Time: 0.358965 seconds (Warm-up)
0.452625 seconds (Sampling)
0.81159 seconds (Total)

Compare the Outputs

lm

```
freq_model_gender <- lm(LS ~ 1 + N + widowed + gender, data = bayes_data)
summary(freq_model_gender)
```

```
## 
## Call:
## lm(formula = LS ~ 1 + N + widowed + gender, data = bayes_data)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3.3425 -0.4567  0.1540  0.6099  1.9070 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 4.26782   0.03597 118.650 < 2e-16 ***
## N          -0.45847   0.02658 -17.250 < 2e-16 ***
## widowedWidowed -0.16425   0.05142  -3.195  0.00142 ** 
## genderFemale  0.07469   0.03330   2.243  0.02499 *  
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 0.8265 on 2702 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.102, Adjusted R-squared:  0.1011 
## F-statistic: 102.4 on 3 and 2702 DF,  p-value: < 2.2e-16
```

brms

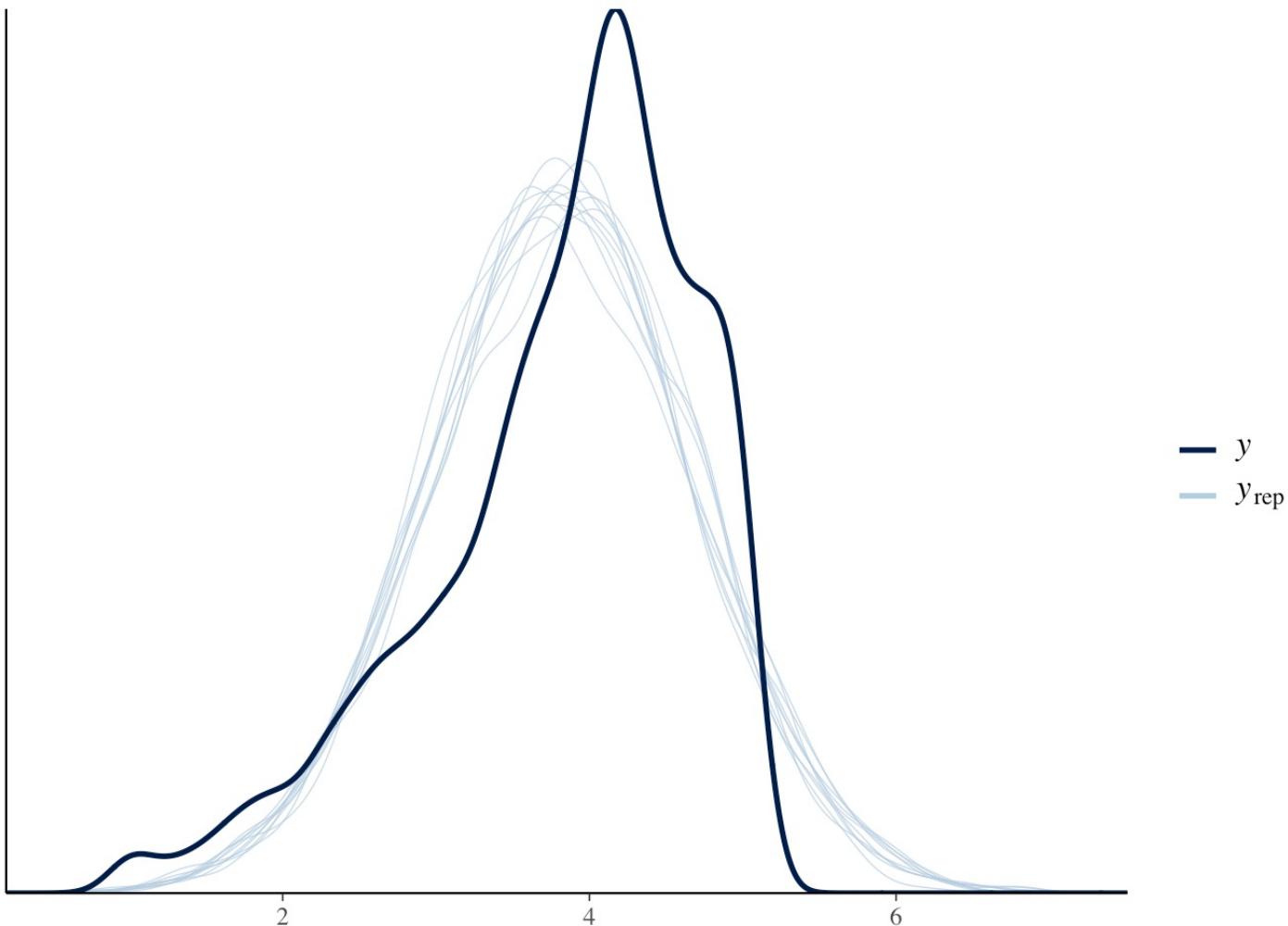
```
summary(brms_model_ND)
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LS ~ 1 + N + widowed + gender
## Data: bayes_data (Number of observations: 2706)
## Draws: 4 chains, each with iter = 6000; warmup = 3000; thin = 1;
##        total post-warmup draws = 12000
##
## Population-Level Effects:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept       4.27     0.04    4.20    4.34 1.00  14449   9034
## N              -0.46     0.03   -0.51   -0.41 1.00  14079   9626
## widowedWidowed -0.16     0.05   -0.26   -0.06 1.00  14129   9449
## genderFemale    0.07     0.03    0.01    0.14 1.00  13411   9716
## 
## Family Specific Parameters:
##             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.83     0.01    0.81    0.85 1.00  14480   9533
## 
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

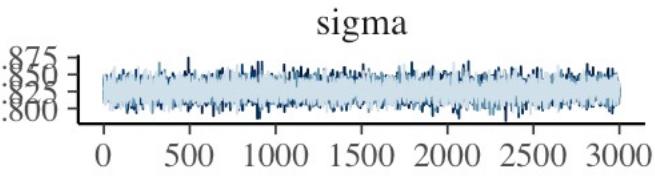
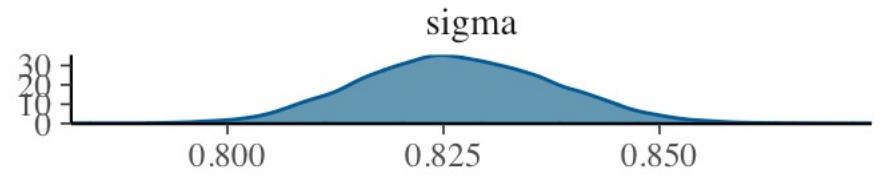
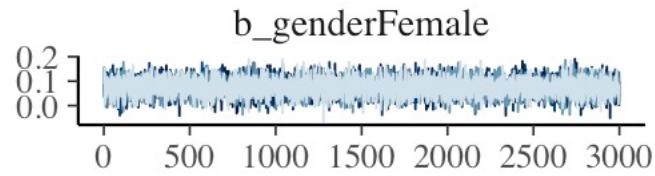
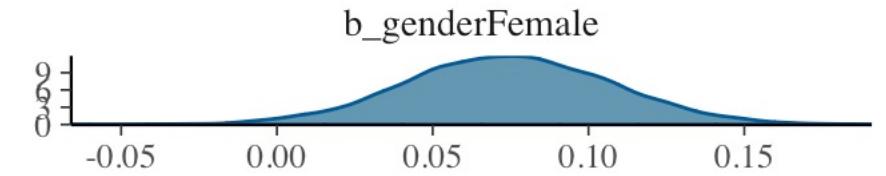
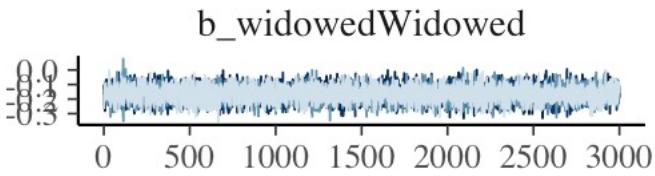
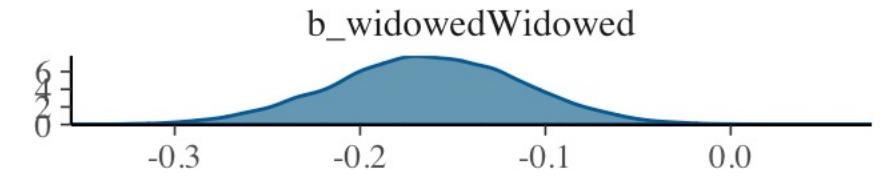
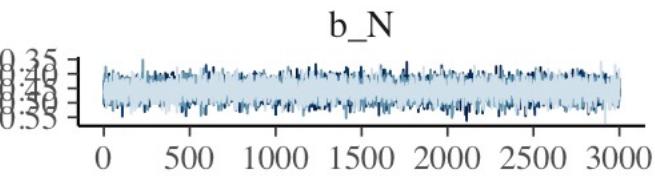
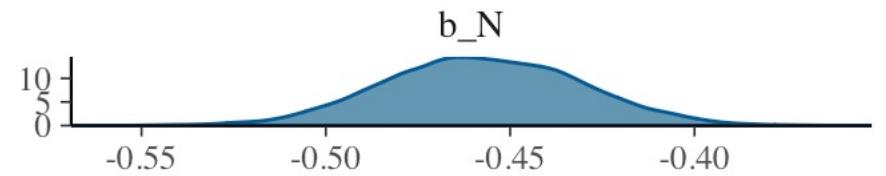
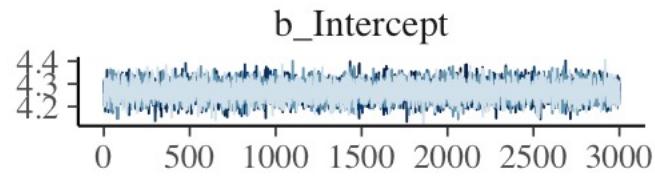
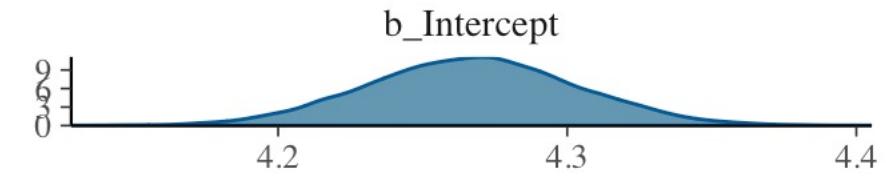
Posterior Predictive Check

```
pp_check(brms_model_ND)
```

```
## Using 10 posterior draws for ppc type 'dens_overlay' by default.
```



```
plot(brms_model_ND)
```

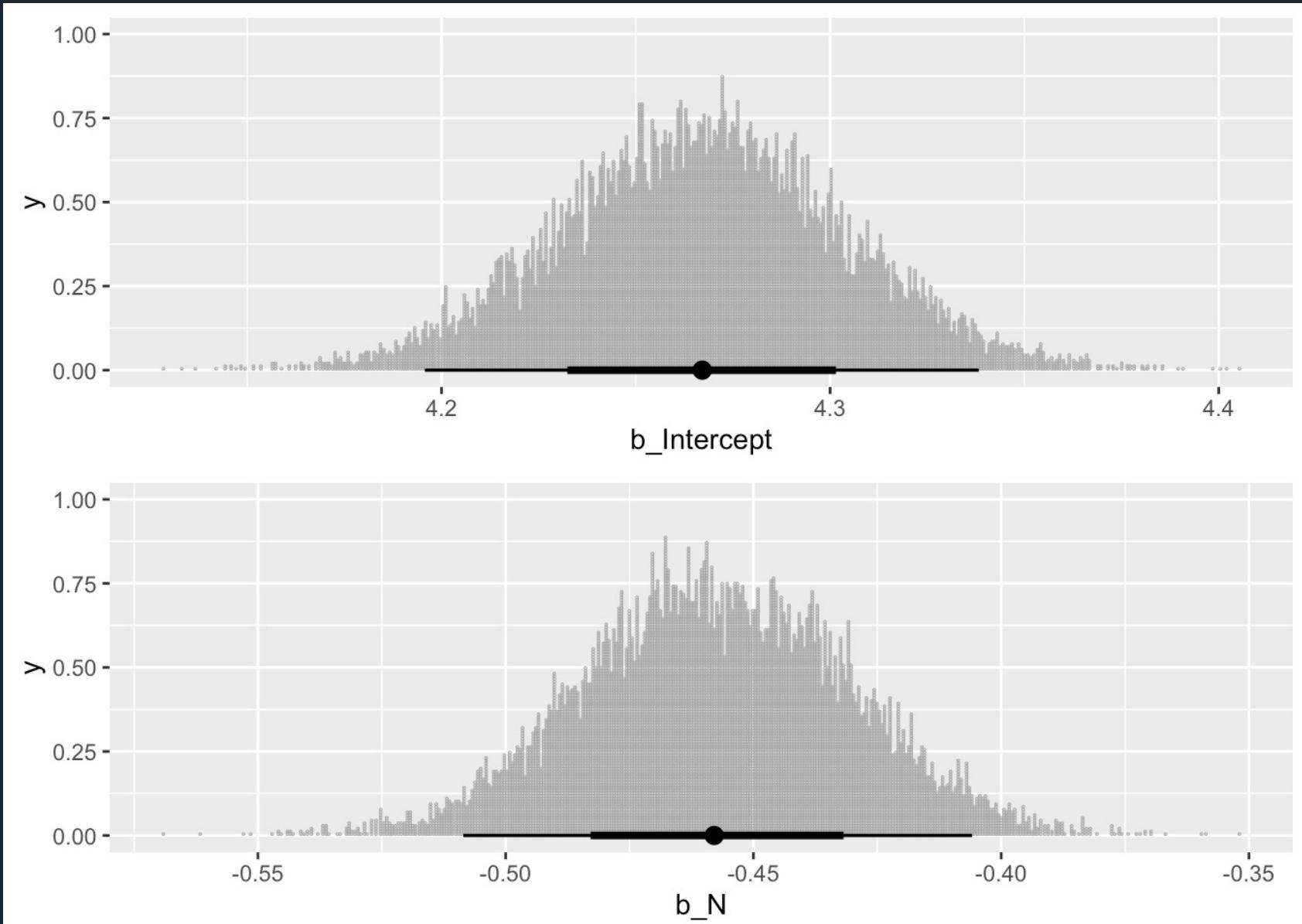


Chain

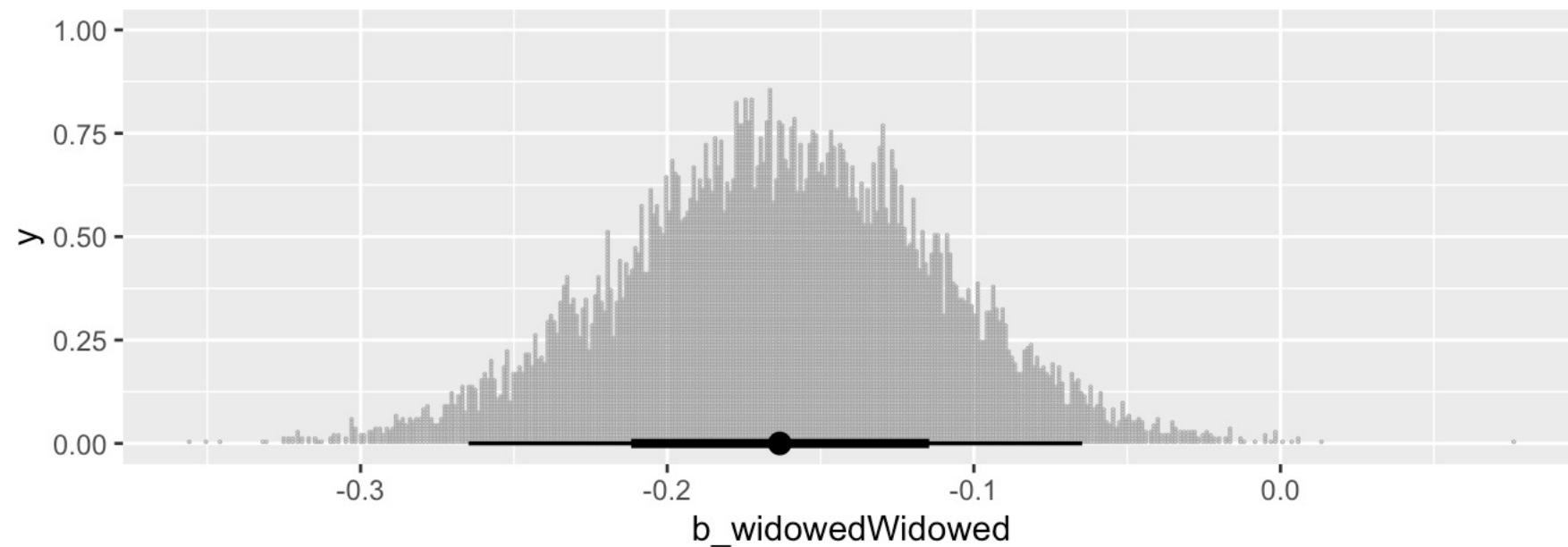
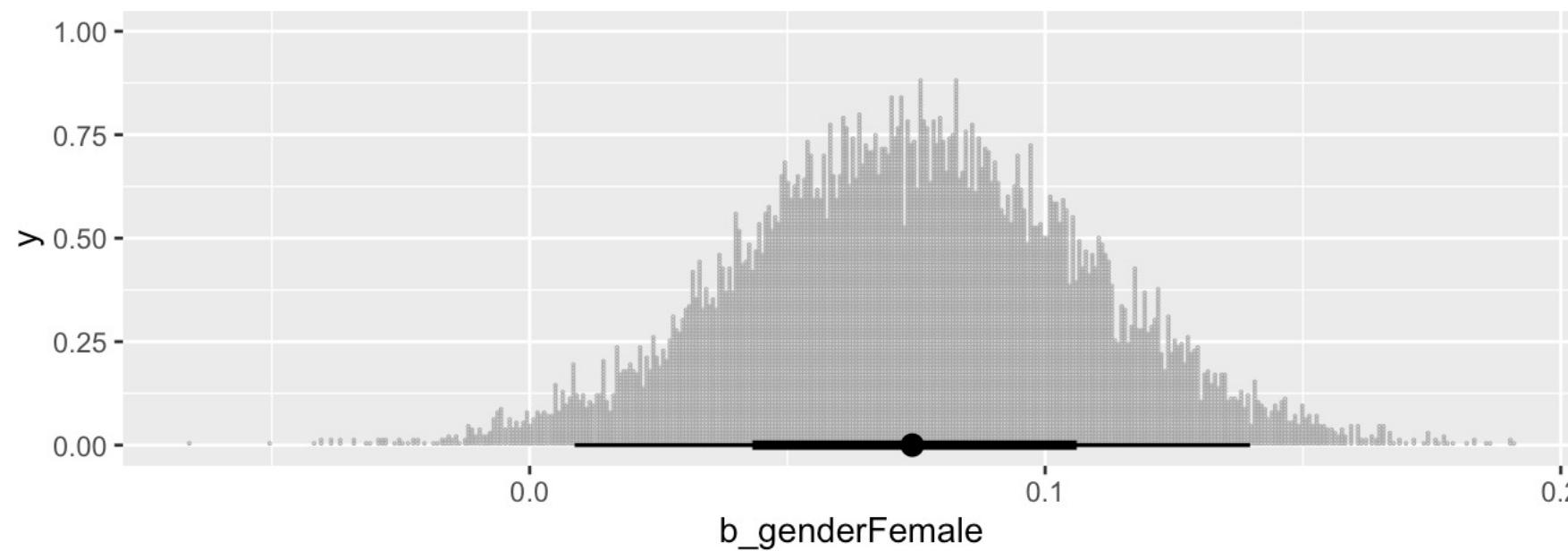
- 1
- 2
- 3
- 4

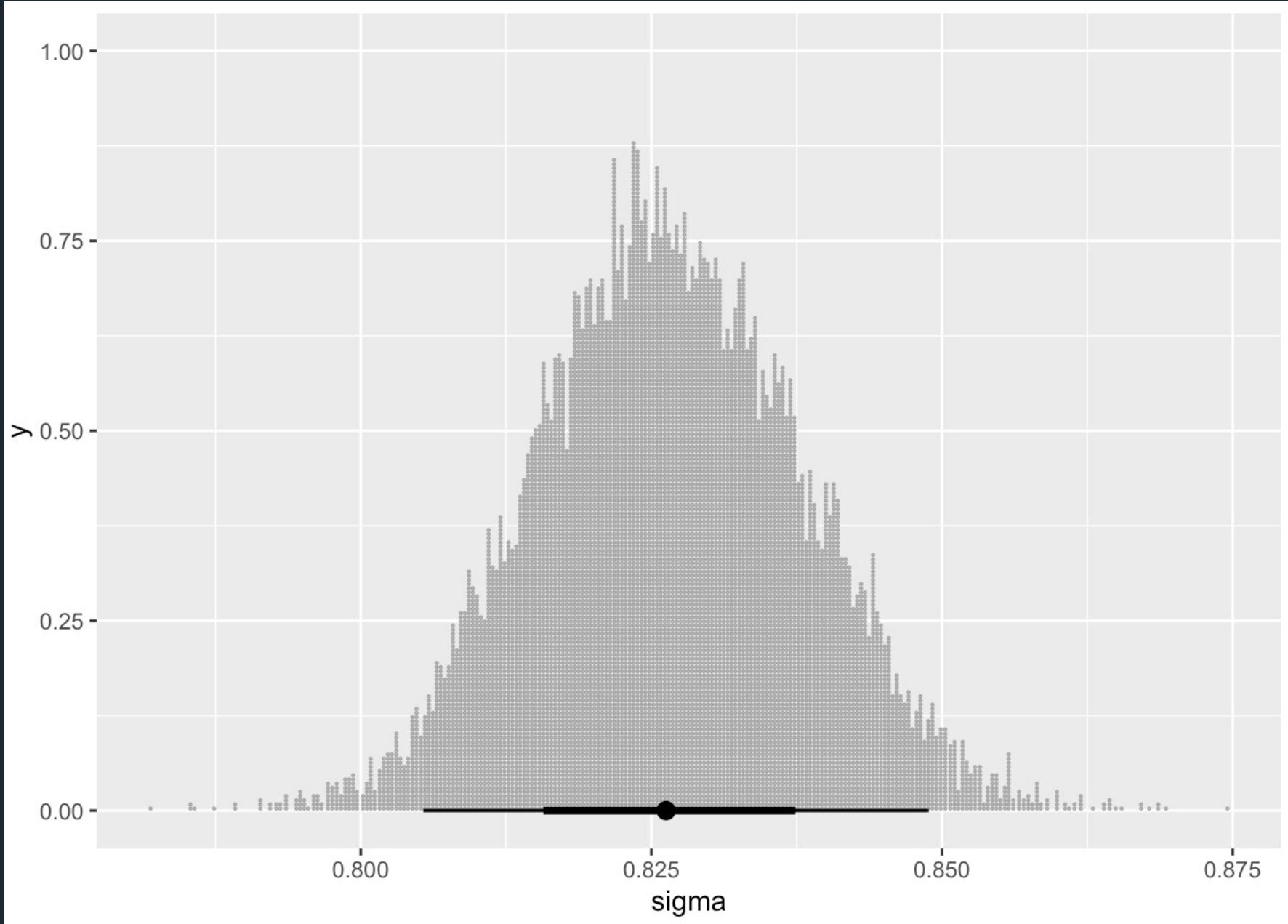
*Checking the
Sampling*

*Distribution of
Estimates: Fixed
Effects*



*Distribution of
Estimates: Fixed
Effects*

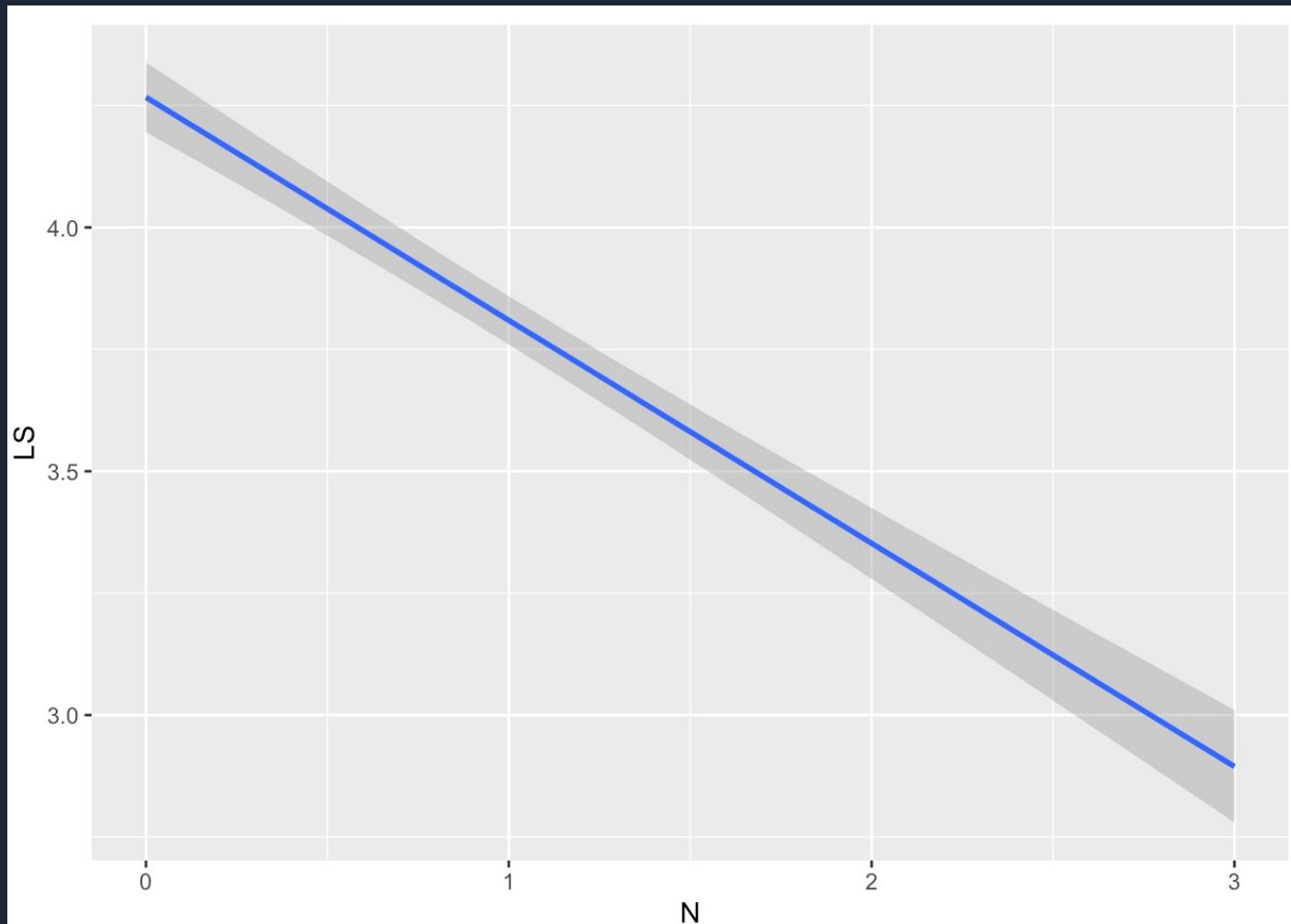




*Distribution of
Estimates:
Family-specific
effects (σ)*

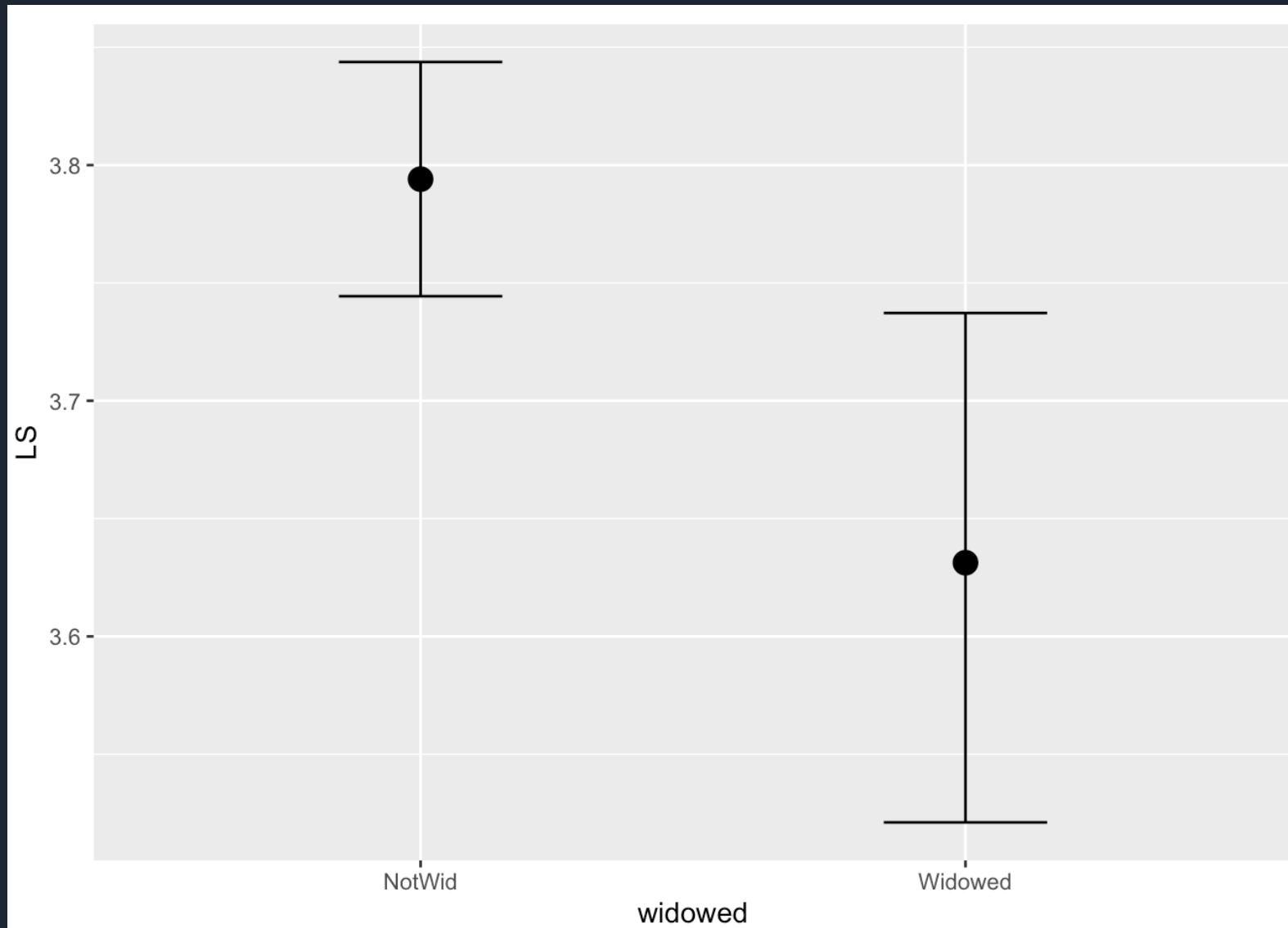
Plot of brms

LS as a function of N scores



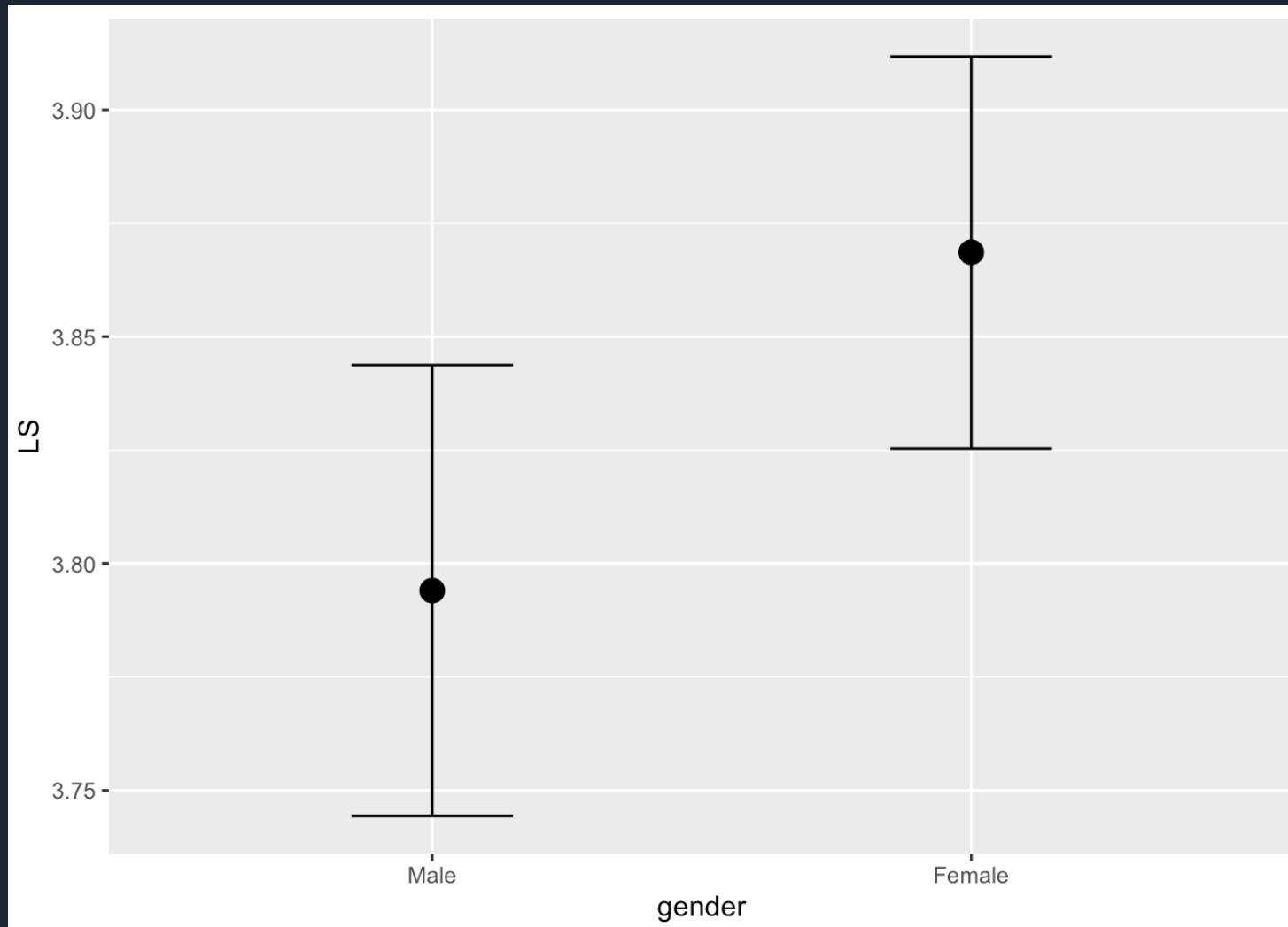
Plot of brms

LS as a function of Being Widowed



Plot of brms

LS as a function of Gender



What about the interaction?

```
# interaction
brms_model.int_ND <-
  brm(family = gaussian,
       LS ~ 1 + N + widowed + gender + widowed:gender,
       prior = c(prior(normal(4, .5), class = Intercept),
                 prior(normal(0, .5), class = b),
                 prior(exponential(1), class = sigma)),
       chains = 4, cores = 4, iter = 6000, warmup = 3000,
       data = bayes_data,
       file = "brms_model.int_ND")
```

```
summary(brms_model.int_ND)

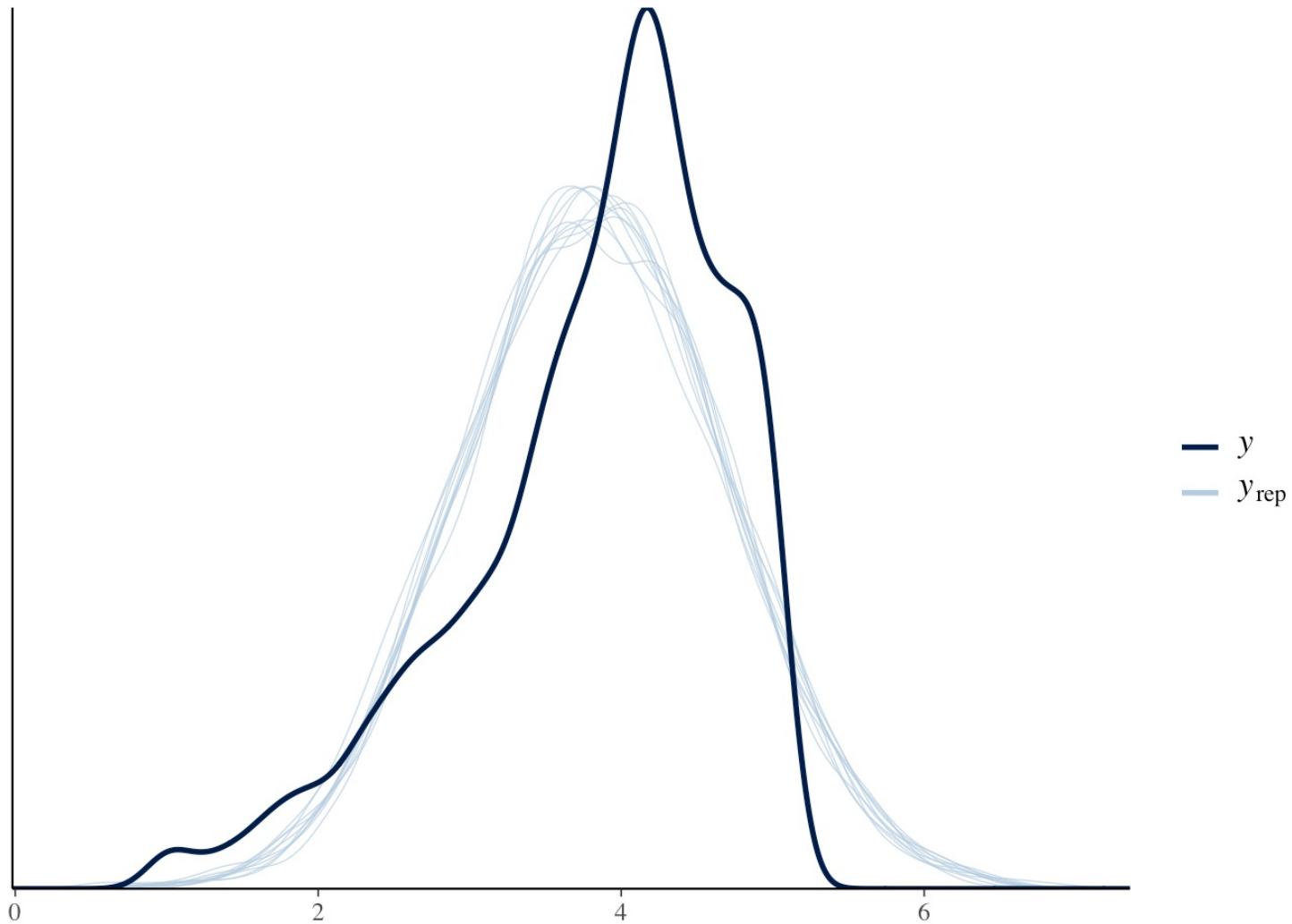
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LS ~ 1 + N + widowed + gender + widowed:gender
## Data: bayes_data (Number of observations: 2706)
## Draws: 4 chains, each with iter = 6000; warmup = 3000; thin = 1;
##        total post-warmup draws = 12000
##
## Population-Level Effects:
##                               Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS
## Intercept                  4.27     0.04    4.20    4.34 1.00   14841
## N                          -0.46     0.03   -0.51   -0.40 1.00   14582
## widowedWidowed             -0.19     0.13   -0.43    0.05 1.00    9755
## genderFemale                0.07     0.03    0.00    0.14 1.00   14117
## widowedWidowed:genderFemale  0.03     0.14   -0.24    0.30 1.00    9555
##                               Tail_ESS
## Intercept                   8457
## N                           8053
## widowedWidowed              7551
## genderFemale                 8471
## widowedWidowed:genderFemale  8119
##
## Family Specific Parameters:
##                               Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma          0.83      0.01     0.81     0.85 1.00   14305     8353
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

brms output interaction

Elapsed Time: 0.751844 seconds (Warm-up)
0.860198 seconds (Sampling)
1.61204 seconds (Total)

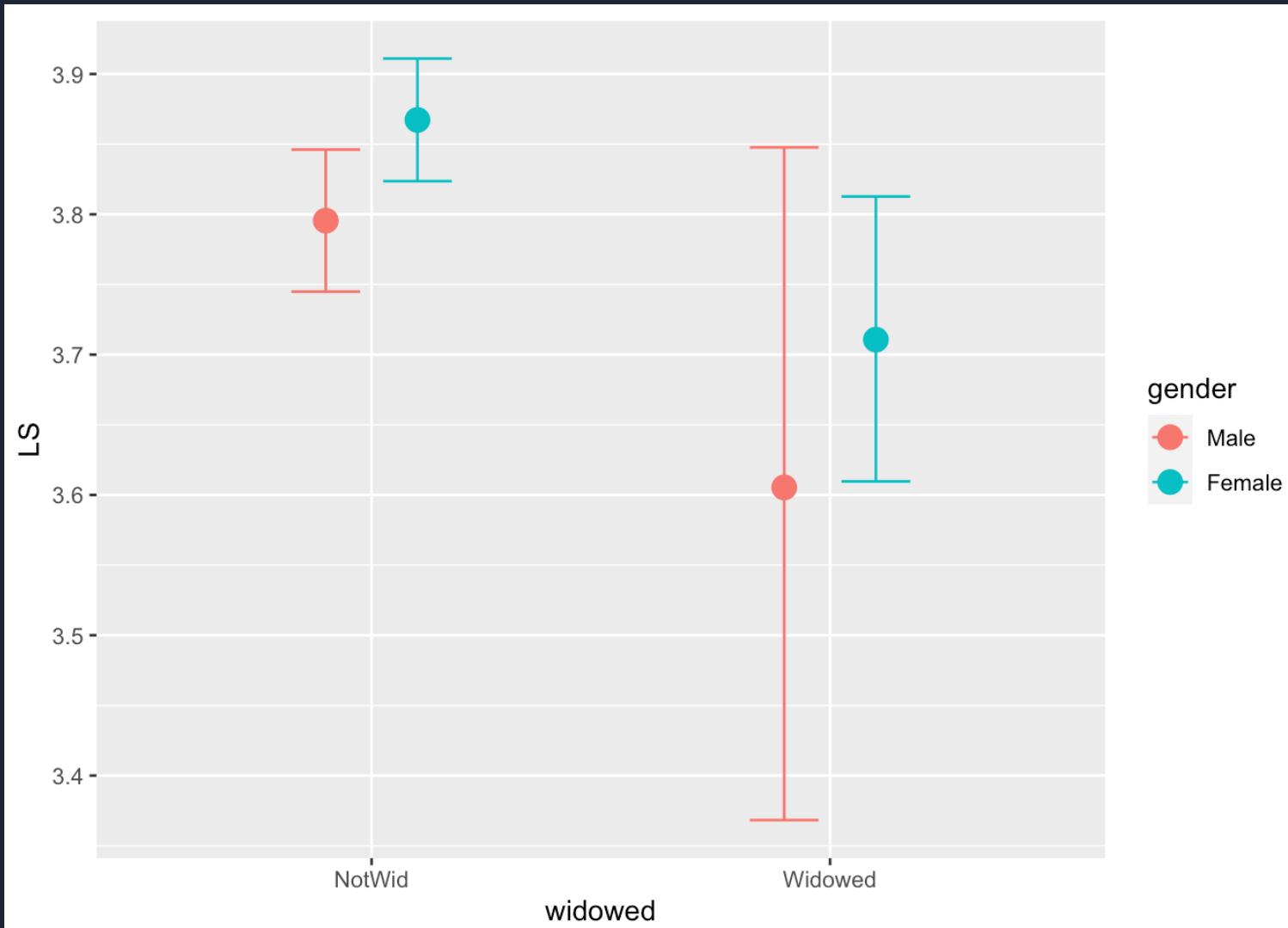
Posterior Predictive Check

```
pp_check(brms_model.int_ND)  
  
## Using 10 posterior draws for ppc type 'dens_overlay' by default.
```



Plot of brms

LS as a function of Widowhood and Gender

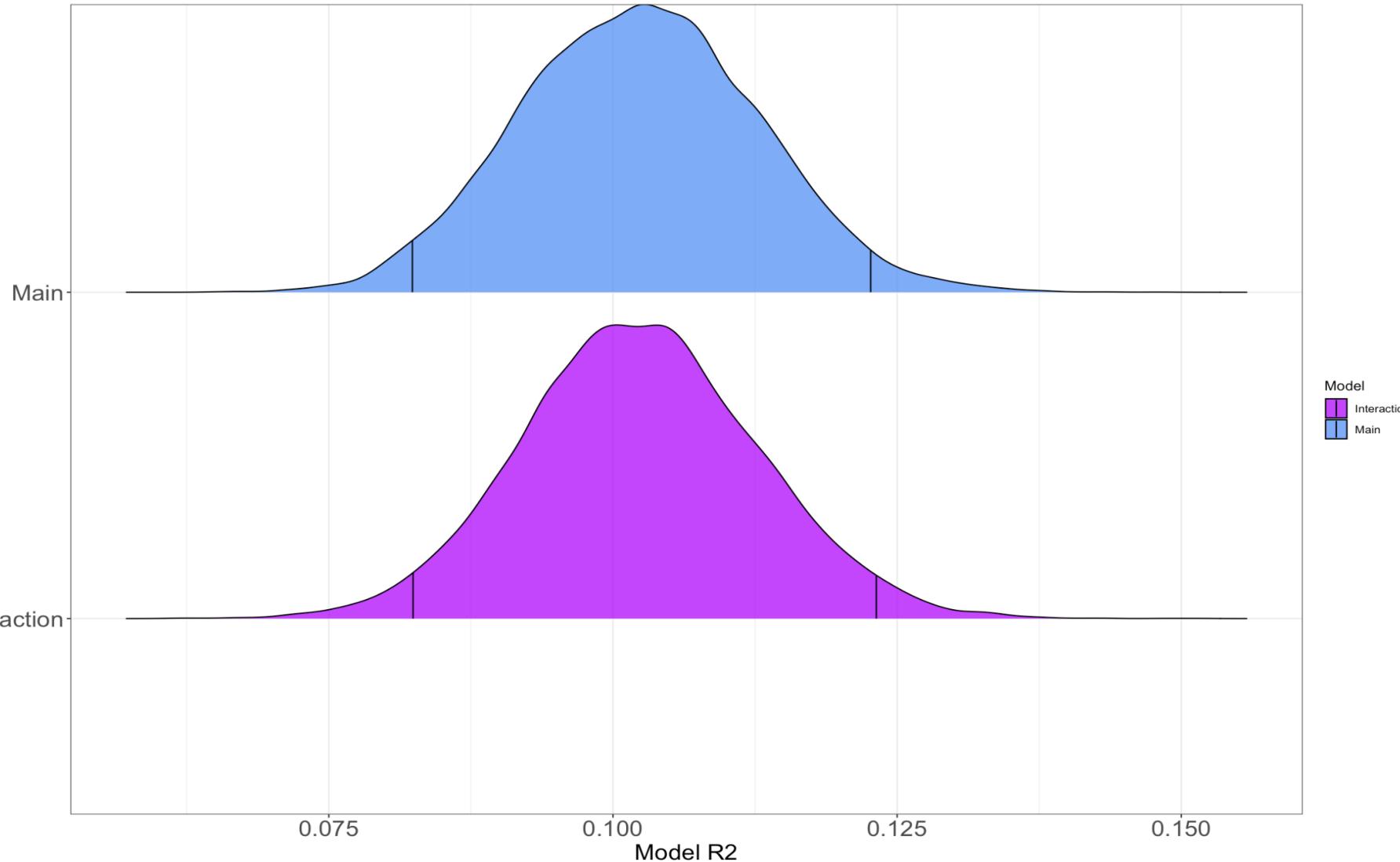


R2 Distributions

Main: 0.1023189

Int: 0.1023718

R2 Distributions for Main Effects Model and Interaction Model



Model Comparisons

```
main_comp <- add_criterion(brms_model_ND, c("loo", "waic"))
int_comp <- add_criterion(brms_model.int_ND, c("loo", "waic"))
comp <- loo_compare(main_comp, int_comp, criterion = "waic")
comp_loo <- loo_compare(main_comp, int_comp, criterion = "loo")
print(comp_loo)

##           elpd_diff se_diff
## main_comp   0.0      0.0
## int_comp  -0.8      0.2

(mw_gsoep <- model_weights(brms_model_ND, brms_model.int_ND))

##       brms_model_ND brms_model.int_ND
## 9.999247e-01    7.529158e-05

mw_gsoep[1] / mw_gsoep[2]

## brms_model_ND
## 13280.7

model_weights(brms_model_ND, brms_model.int_ND, weights = "waic")

##       brms_model_ND brms_model.int_ND
## 0.699141      0.300859
```

Pros & Cons of Bayesian Approaches

Pros

- More flexible ways to specify your model
- Fewer worries of lack of convergence
- Fewer & more flexible assumptions
- You interpret estimates and statistics how you likely naturally think about them (e.g., CIs)
- Less reliance on “significance”
- Requires more upfront thought about your models, should translate into a better understanding of them
- Type I/II errors, in their most technical definitions, do not exist

Cons

- Can be quite time-consuming (time is ~directly proportional to complexity of model and amount of data you have)
- Greater computational resources are likely necessary for more complex models
- Will exclude rows in your data frame with NAs (i.e., missing data; `lm()` does this too, though)
- Still has some misconceptions surrounding it

Helpful Resources

- Kruschke's Bayesian new statistics: <https://rdcu.be/bRUvW>
- Richard McElreath's *Statistical Rethinking*: <https://xcelab.net/rm/statistical-rethinking/>
- Solomon Kurz's version/translation of *Statistical Rethinking* using R code: <https://bookdown.org/content/4857/>