Bayes approach

Saneesh

2022-12-02

Generalized linear models

```
library(brms)
```

```
## Loading required package: Rcpp
```

```
## Loading 'brms' package (version 2.18.0). Useful instructions
## can be found by typing help('brms'). A more detailed introduction
## to the package is available through vignette('brms_overview').
```

```
##
## Attaching package: 'brms'
```

```
## The following object is masked from 'package:stats':
##
##     ar
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
```

```
## v ggplot2 3.4.0      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(report)
cardata <- brm(vs ~ mpg * drat,
               data = mtcars,
               family = bernoulli())
```

```
## Compiling Stan program...
## Start sampling
```

```
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.366 seconds (Warm-up)
## Chain 1:                0.257 seconds (Sampling)
## Chain 1:                0.623 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 0 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 2: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 2: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 2: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 2: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 2: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 2: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 2: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 0.317 seconds (Warm-up)
## Chain 2:                0.244 seconds (Sampling)
## Chain 2:                0.561 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 0 seconds
```

```
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 3: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 3: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 3: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 3: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 0.311 seconds (Warm-up)
## Chain 3:                0.271 seconds (Sampling)
## Chain 3:                0.582 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 4).
## Chain 4:
## Chain 4: Gradient evaluation took 0 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 4: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 4: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 4: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 4: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 4: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 0.309 seconds (Warm-up)
## Chain 4:                0.278 seconds (Sampling)
## Chain 4:                0.587 seconds (Total)
## Chain 4:
```

```
summary(cardata)
```

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: vs ~ mpg * drat
##    Data: mtcars (Number of observations: 32)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
```

```
##          total post-warmup draws = 4000
##
## Population-Level Effects:
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept   -36.54     23.37   -87.11     6.24 1.00      751     1139
## mpg           1.99      1.20    -0.22     4.53 1.00      758     1154
## drat          5.94      5.78    -5.17    17.64 1.00      800     1151
## mpg:drat     -0.34      0.28    -0.91     0.21 1.00      754     1120
##
## 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).
```

```
report(cardata)
```

```
## Warning: Response residuals not available to calculate mean square error. (R)MSE
##   is probably not reliable.

## Start sampling

##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.063 seconds (Warm-up)
## Chain 1:                0.069 seconds (Sampling)
## Chain 1:                0.132 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 0 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
```

```
## Chain 2: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 2: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 2: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 2: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 2: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 2: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 2: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 2: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 0.072 seconds (Warm-up)
## Chain 2:                0.072 seconds (Sampling)
## Chain 2:                0.144 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 0 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 3: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 3: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 3: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 3: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 0.061 seconds (Warm-up)
## Chain 3:                0.065 seconds (Sampling)
## Chain 3:                0.126 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL '63230c89284e82c150c4ad8a7dc4eb7c' NOW (CHAIN 4).
## Chain 4:
## Chain 4: Gradient evaluation took 0 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 4: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 4: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 4: Iteration:  600 / 2000 [ 30%]  (Warmup)
```

```
## Chain 4: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 4: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 0.062 seconds (Warm-up)
## Chain 4:                0.069 seconds (Sampling)
## Chain 4:                0.131 seconds (Total)
## Chain 4:
```

```
## Warning: Response residuals not available to calculate mean square error. (R)MSE
##   is probably not reliable.
```

```
## We fitted a Bayesian logistic model (estimated using MCMC sampling with 4
## chains of 2000 iterations and a warmup of 1000) to predict vs with mpg
## (formula: vs ~ mpg * drat). Priors over parameters were set as student_t
## (location = 0.00, scale = 2.50) distributions. The model's explanatory power is
## substantial (R2 = 0.51, 95% CI [0.40, 0.57]).  Within this model:
##
##   - The effect of b Intercept (Median = -35.37, 95% CI [-87.11, 6.24]) has a
## 95.17% probability of being negative (< 0), 95.15% of being significant (<
## -0.09), and 94.92% of being large (< -0.54). The estimation successfully
## converged (Rhat = 1.001) but the indices are unreliable (ESS = 798)
##   - The effect of b mpg (Median = 1.95, 95% CI [-0.22, 4.53]) has a 96.20%
## probability of being positive (> 0), 95.40% of being significant (> 0.09), and
## 88.52% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.001) but the indices are unreliable (ESS = 753)
##   - The effect of b drat (Median = 5.84, 95% CI [-5.17, 17.64]) has a 84.50%
## probability of being positive (> 0), 84.03% of being significant (> 0.09), and
## 82.15% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.002) but the indices are unreliable (ESS = 756)
##   - The interaction effect of drat on b mpg (Median = -0.34, 95% CI [-0.91,
## 0.21]) has a 88.55% probability of being negative (< 0), 81.08% of being
## significant (< -0.09), and 23.08% of being large (< -0.54). The estimation
## successfully converged (Rhat = 1.001) but the indices are unreliable (ESS =
## 750)
##
## Following the Sequential Effect eXistence and sIgnificance Testing (SEXIT)
## framework, we report the median of the posterior distribution and its 95% CI
## (Highest Density Interval), along the probability of direction (pd), the
## probability of significance and the probability of being large. The thresholds
## beyond which the effect is considered as significant (i.e., non-negligible) and
## large are |0.09| and |0.54|. Convergence and stability of the Bayesian sampling
## has been assessed using R-hat, which should be below 1.01 (Vehtari et al.,
## 2019), and Effective Sample Size (ESS), which should be greater than 1000
## (Burkner, 2017)., We fitted a Bayesian logistic model (estimated using MCMC
## sampling with 4 chains of 2000 iterations and a warmup of 1000) to predict vs
## with drat (formula: vs ~ mpg * drat). Priors over parameters were set as
## uniform (location = , scale = ) distributions. The model's explanatory power is
## substantial (R2 = 0.51, 95% CI [0.40, 0.57]).  Within this model:
```

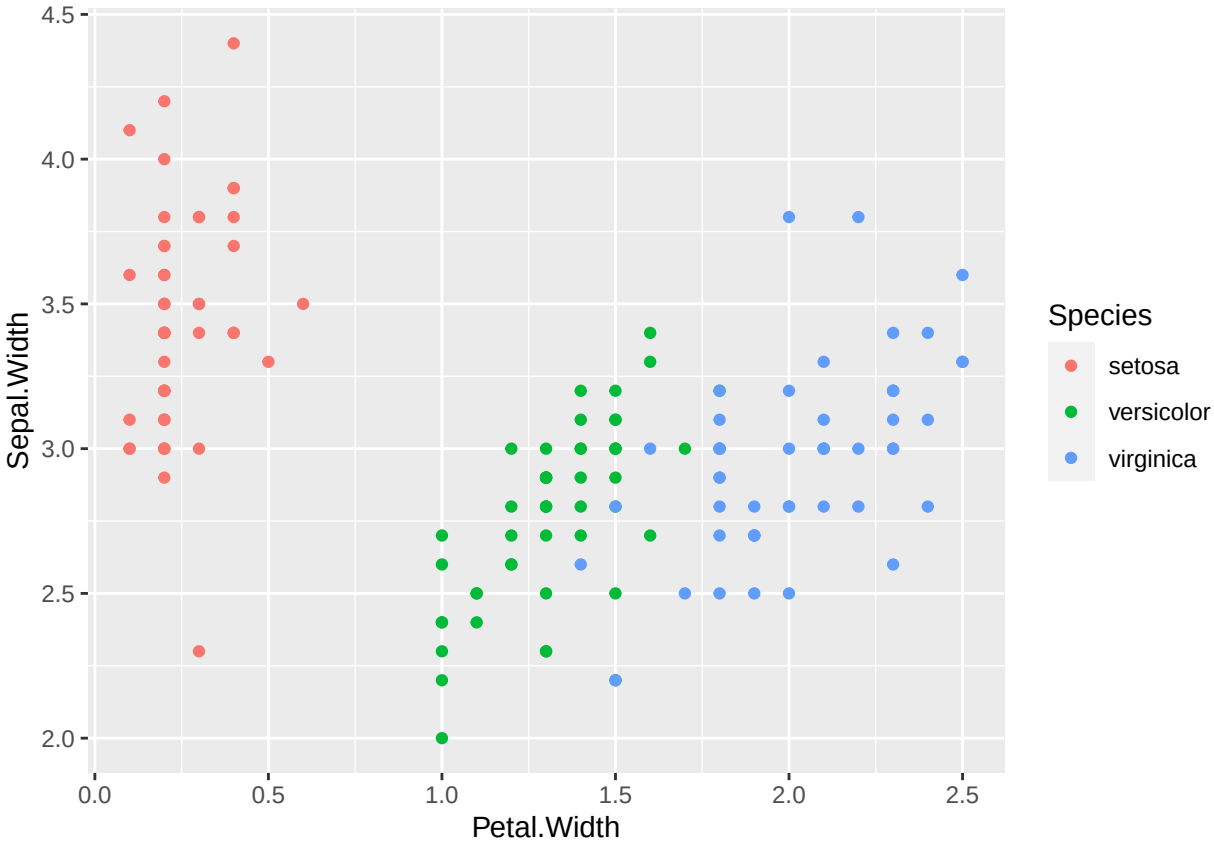
```
##
##   - The effect of b Intercept (Median = -35.37, 95% CI [-87.11, 6.24]) has a
## 95.17% probability of being negative (< 0), 95.15% of being significant (<
## -0.09), and 94.92% of being large (< -0.54). The estimation successfully
## converged (Rhat = 1.001) but the indices are unreliable (ESS = 798)
##   - The effect of b mpg (Median = 1.95, 95% CI [-0.22, 4.53]) has a 96.20%
## probability of being positive (> 0), 95.40% of being significant (> 0.09), and
## 88.52% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.001) but the indices are unreliable (ESS = 753)
##   - The effect of b drat (Median = 5.84, 95% CI [-5.17, 17.64]) has a 84.50%
## probability of being positive (> 0), 84.03% of being significant (> 0.09), and
## 82.15% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.002) but the indices are unreliable (ESS = 756)
##   - The interaction effect of drat on b mpg (Median = -0.34, 95% CI [-0.91,
## 0.21]) has a 88.55% probability of being negative (< 0), 81.08% of being
## significant (< -0.09), and 23.08% of being large (< -0.54). The estimation
## successfully converged (Rhat = 1.001) but the indices are unreliable (ESS =
## 750)
##
## Following the Sequential Effect eXistence and sIgnificance Testing (SEXIT)
## framework, we report the median of the posterior distribution and its 95% CI
## (Highest Density Interval), along the probability of direction (pd), the
## probability of significance and the probability of being large. The thresholds
## beyond which the effect is considered as significant (i.e., non-negligible) and
## large are |0.09| and |0.54|. Convergence and stability of the Bayesian sampling
## has been assessed using R-hat, which should be below 1.01 (Vehtari et al.,
## 2019), and Effective Sample Size (ESS), which should be greater than 1000
## (Burkner, 2017)., We fitted a Bayesian logistic model (estimated using MCMC
## sampling with 4 chains of 2000 iterations and a warmup of 1000) to predict vs
## with mpg (formula: vs ~ mpg * drat). Priors over parameters were set as uniform
## (location = , scale = ) distributions. The model's explanatory power is
## substantial (R2 = 0.51, 95% CI [0.40, 0.57]).  Within this model:
##
##   - The effect of b Intercept (Median = -35.37, 95% CI [-87.11, 6.24]) has a
## 95.17% probability of being negative (< 0), 95.15% of being significant (<
## -0.09), and 94.92% of being large (< -0.54). The estimation successfully
## converged (Rhat = 1.001) but the indices are unreliable (ESS = 798)
##   - The effect of b mpg (Median = 1.95, 95% CI [-0.22, 4.53]) has a 96.20%
## probability of being positive (> 0), 95.40% of being significant (> 0.09), and
## 88.52% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.001) but the indices are unreliable (ESS = 753)
##   - The effect of b drat (Median = 5.84, 95% CI [-5.17, 17.64]) has a 84.50%
## probability of being positive (> 0), 84.03% of being significant (> 0.09), and
## 82.15% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.002) but the indices are unreliable (ESS = 756)
##   - The interaction effect of drat on b mpg (Median = -0.34, 95% CI [-0.91,
## 0.21]) has a 88.55% probability of being negative (< 0), 81.08% of being
## significant (< -0.09), and 23.08% of being large (< -0.54). The estimation
## successfully converged (Rhat = 1.001) but the indices are unreliable (ESS =
## 750)
##
## Following the Sequential Effect eXistence and sIgnificance Testing (SEXIT)
## framework, we report the median of the posterior distribution and its 95% CI
## (Highest Density Interval), along the probability of direction (pd), the
```

```
## probability of significance and the probability of being large. The thresholds
## beyond which the effect is considered as significant (i.e., non-negligible) and
## large are |0.09| and |0.54|. Convergence and stability of the Bayesian sampling
## has been assessed using R-hat, which should be below 1.01 (Vehtari et al.,
## 2019), and Effective Sample Size (ESS), which should be greater than 1000
## (Burkner, 2017). and We fitted a Bayesian logistic model (estimated using MCMC
## sampling with 4 chains of 2000 iterations and a warmup of 1000) to predict vs
## with drat (formula: vs ~ mpg * drat). Priors over parameters were set as
## uniform (location = , scale = ) distributions. The model's explanatory power is
## substantial (R2 = 0.51, 95% CI [0.40, 0.57]).  Within this model:
##
##   - The effect of b Intercept (Median = -35.37, 95% CI [-87.11, 6.24]) has a
## 95.17% probability of being negative (< 0), 95.15% of being significant (<
## -0.09), and 94.92% of being large (< -0.54). The estimation successfully
## converged (Rhat = 1.001) but the indices are unreliable (ESS = 798)
##   - The effect of b mpg (Median = 1.95, 95% CI [-0.22, 4.53]) has a 96.20%
## probability of being positive (> 0), 95.40% of being significant (> 0.09), and
## 88.52% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.001) but the indices are unreliable (ESS = 753)
##   - The effect of b drat (Median = 5.84, 95% CI [-5.17, 17.64]) has a 84.50%
## probability of being positive (> 0), 84.03% of being significant (> 0.09), and
## 82.15% of being large (> 0.54). The estimation successfully converged (Rhat =
## 1.002) but the indices are unreliable (ESS = 756)
##   - The interaction effect of drat on b mpg (Median = -0.34, 95% CI [-0.91,
## 0.21]) has a 88.55% probability of being negative (< 0), 81.08% of being
## significant (< -0.09), and 23.08% of being large (< -0.54). The estimation
## successfully converged (Rhat = 1.001) but the indices are unreliable (ESS =
## 750)
##
## Following the Sequential Effect eXistence and sIgnificance Testing (SEXIT)
## framework, we report the median of the posterior distribution and its 95% CI
## (Highest Density Interval), along the probability of direction (pd), the
## probability of significance and the probability of being large. The thresholds
## beyond which the effect is considered as significant (i.e., non-negligible) and
## large are |0.09| and |0.54|. Convergence and stability of the Bayesian sampling
## has been assessed using R-hat, which should be below 1.01 (Vehtari et al.,
## 2019), and Effective Sample Size (ESS), which should be greater than 1000
## (Burkner, 2017).
```

Mixed model with random slopes.
Question
>Is there a general relationship between petal and sepal width? and how it differs by species?

```
library(brms)
library(tidyverse)
library(report)

ggplot(iris, aes(Petal.Width, Sepal.Width, colour= Species))+
  geom_point()
```


4.5
4.0
3.5
3.0
2.5
2.0
Sepal.Width
0.0
0.5
1.0
1.5
2.0
2.5
Petal.Width
Species
setosa
versicolor
virginica