

Applied Bayesian modeling - HW3, V1

Álvaro J. Castro Rivadeneira

September 22, 2022

I am posting the HW-V1 to have questions to work on in class on 9/22/after watching module 6. Note that there may be additional questions added later.

Exercise 1: Fit a Bayesian model using brm and check and interpret the output

We continue with IQ data, as introduced in HW2. For this HW, data set iq_scores.csv contains 10 iq-scores, sampled from a town in let's say, your favorite country.

```
iqs <- read_csv("iq_scores.csv") %>%
  mutate(y = iq_score)

## Rows: 10 Columns: 1
## -- Column specification -----
## Delimiter: ","
## dbl (1): iq_score
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Use brm to fit the following model to the IQ data:

$$\begin{aligned} y_i | \theta_i, \sigma^2 &\sim N(\theta_i, \sigma^2) (\text{independent}), \text{ for } i = 1, 2, \dots, n; \\ \mu &\sim N(100, 15^2); \\ \sigma &\sim \text{use the brm-default.} \end{aligned}$$

Bayesian regression

```
mu_prior <- set_prior("normal(100,15)", class = "Intercept")

iq_fit <- brm(iq_score ~ 1, data = iqs,
  prior = c(mu_prior),
  file = "output/exercise1",
  chains = 4, iter = 1000, warmup = 500,
  cores = getOption("mc.cores", 4))
```

Quick summary overview

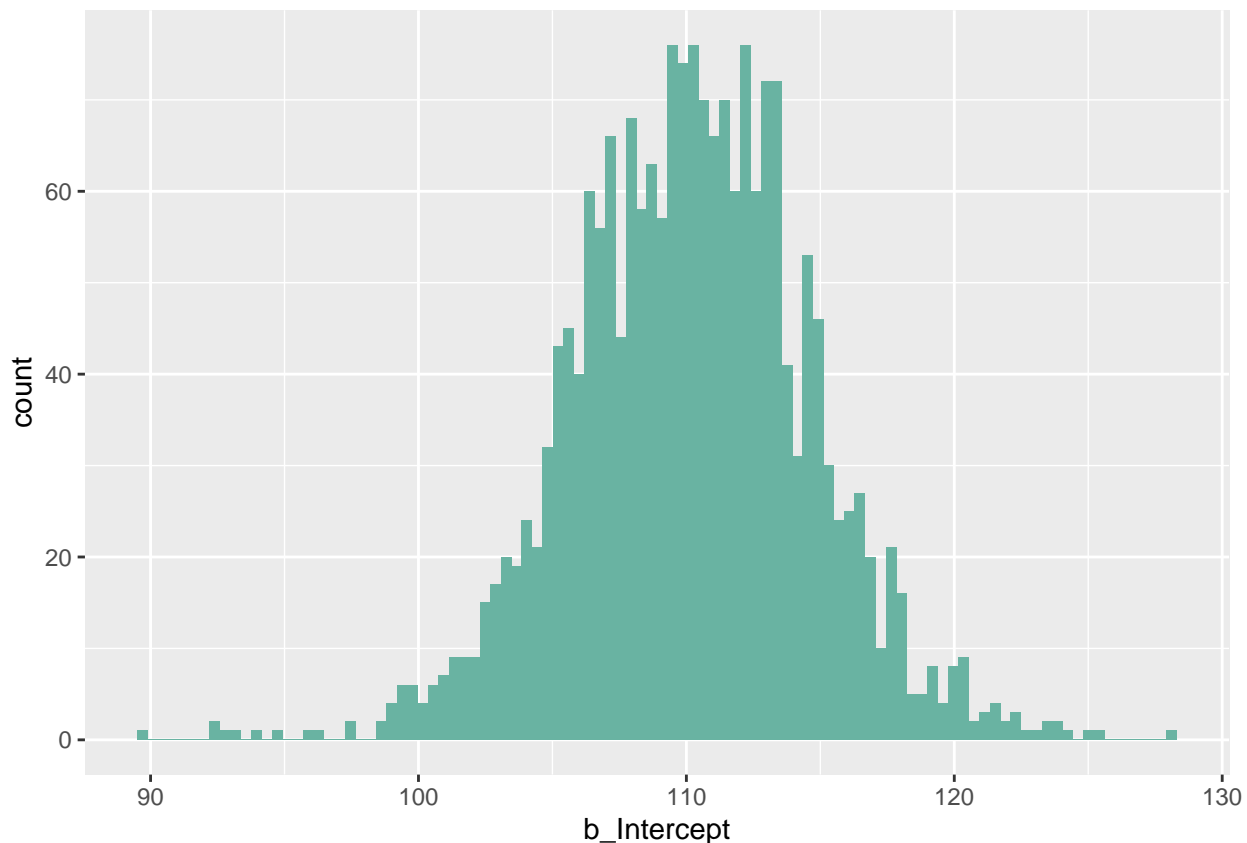
```
summary(iq.fit)
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: iq_score ~ 1
## Data: iqs (Number of observations: 10)
## Draws: 4 chains, each with iter = 1000; warmup = 500; thin = 1;
## total post-warmup draws = 2000
##
## Population-Level Effects:
##      Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept  110.17      4.47  101.34  119.18 1.00    1027    989
##
## Family Specific Parameters:
##      Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma    14.29      3.71   9.18  23.39 1.00    1056    885
##
## 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).
```

Then answer the following questions:

(i) Plot a histogram of the posterior samples of μ and report a posterior point estimate and 80% CI for μ .

```
post.iq.mean <- as_draws_df(iq.fit, variable = c("b_Intercept"))
post.iq.mean %>%
  ggplot(aes(x = b_Intercept)) +
  geom_histogram(bins=100, fill = "#69b3a2")
```



```
posterior_summary(iq.fit, probs = c(.1, .9), variable = "b_Intercept")
```

```
##           Estimate Est.Error      Q10      Q90
## b_Intercept 110.1703   4.470508 104.8133 115.5434
```

(ii) Plot a histogram of the posterior samples of σ and report a posterior point estimate and 80% CI for σ .

(iii) Can you report a posterior point estimate and 80% CI for μ/σ ? If yes, do so. If not, why not?

Note that the prior for μ can be specified with an additional argument in `brm` (as illustrated in optional material in module 5), as follows:

```
#mu_prior <- set_prior("normal(100,15)", class = "Intercept")
#fit <- brm(y ~ 1, prior = c(mu_prior),
#           your other usual arguments)
```

Exercise 2: Compare and contrast the MCMC diagnostics of two different model fits

Continue with the IQ data from Q1 (with y = IQ scores) to fit the model as specified below. Present and briefly summarize resulting MCMC diagnostics for μ (traceplots, Rhat, effective sample sizes). Then and comment on whether this model fit can be used for summarizing information regarding μ . If not, why not?

```
#fit_bad <- brm(y ~ 1, data = dat,  
#           chains = 4, iter = 200, cores = getOption("mc.cores", 4),  
#           control = list(adapt_delta = 0.6, max_treedepth = 4)  
#           # these are NOT recommended options, trying to create problems here!  
#           )
```