

Bayesian models and Markov chains

Mohar Sen

6/14/2020

Research topic: Sleep deprivation

Research Question

How does sleep deprivation impact reaction time?

The Study

- measure reaction time on Day 0
- restrict sleep to 3 hours per night
- measure reaction time on Day 3
- measure the change in reaction time

For subject i , let Y_i be the change in reaction time (in ms) after 3 sleep deprived nights. Of course, people react differently to sleep deprivation. It's reasonable to assume that Y_i are Normally distributed around some average m with standard deviation s

Y_i = change in reaction time(ms) for subject i

Assume

Y_i are Normally distributed around some average change in reaction time m with standard deviation s .

$$Y_i \sim N(m, s^2)$$

Prior model for parameter m

Y_i = change in reaction time (ms)

$$Y_i \sim N(m, s^2)$$

$$m = \text{average} Y_i$$

Prior information:

- with normal sleep, average reaction time is ~250 ms
- expect average to increase by ~50 ms
- average is unlikely to decrease & unlikely to increase by more than ~150 ms

Thus, $m \sim N(250, 150^2)$

Also, * $s > 0$ * with normal sleep, s.d. in reaction times is ~30 ms * s is equally likely to be anywhere from 0 to 200 ms

Thus, $s \sim \text{Unif}(0, 200)$

Therefore, $Y_i \sim N(m, s^2)$ $m \sim N(250, 150^2)$ $s \sim \text{Unif}(0, 200)$

```
library(ggplot2)
```

Normal-Normal priors

In the first step of your Bayesian analysis, you'll simulate the following prior models for parameters m and s :

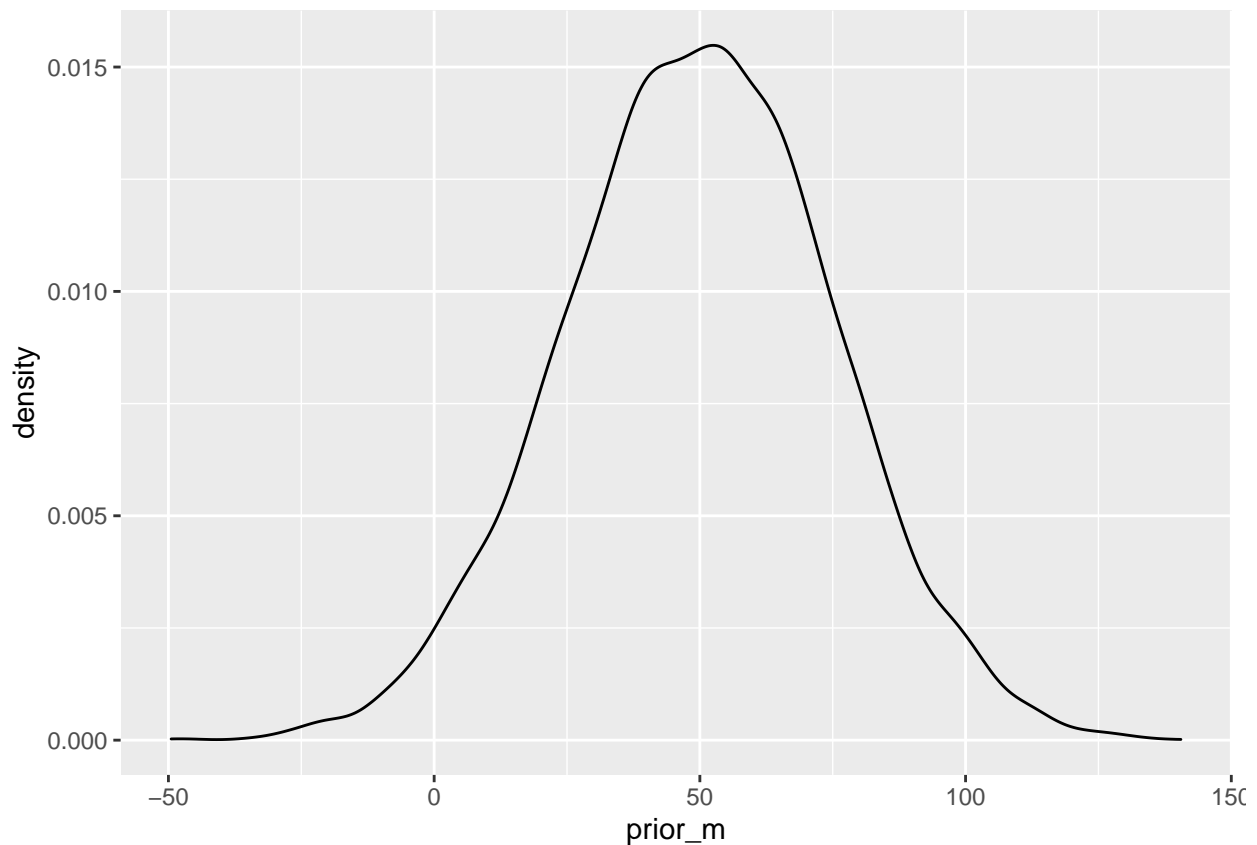
- Use `rnorm(n, mean, sd)` to sample 10,000 draws from the m prior. Assign the output to `prior_m`.
- Use `runif(n, min, max)` to sample 10,000 draws from the s prior. Assign the output to `prior_s`.
- After storing these results in the `samples` data frame, construct a density plot of the `prior_m` samples and a density plot of the `prior_s` samples.

```
# Take 10000 samples from the m prior
prior_m <- rnorm(n=10000, mean=50, sd=25)

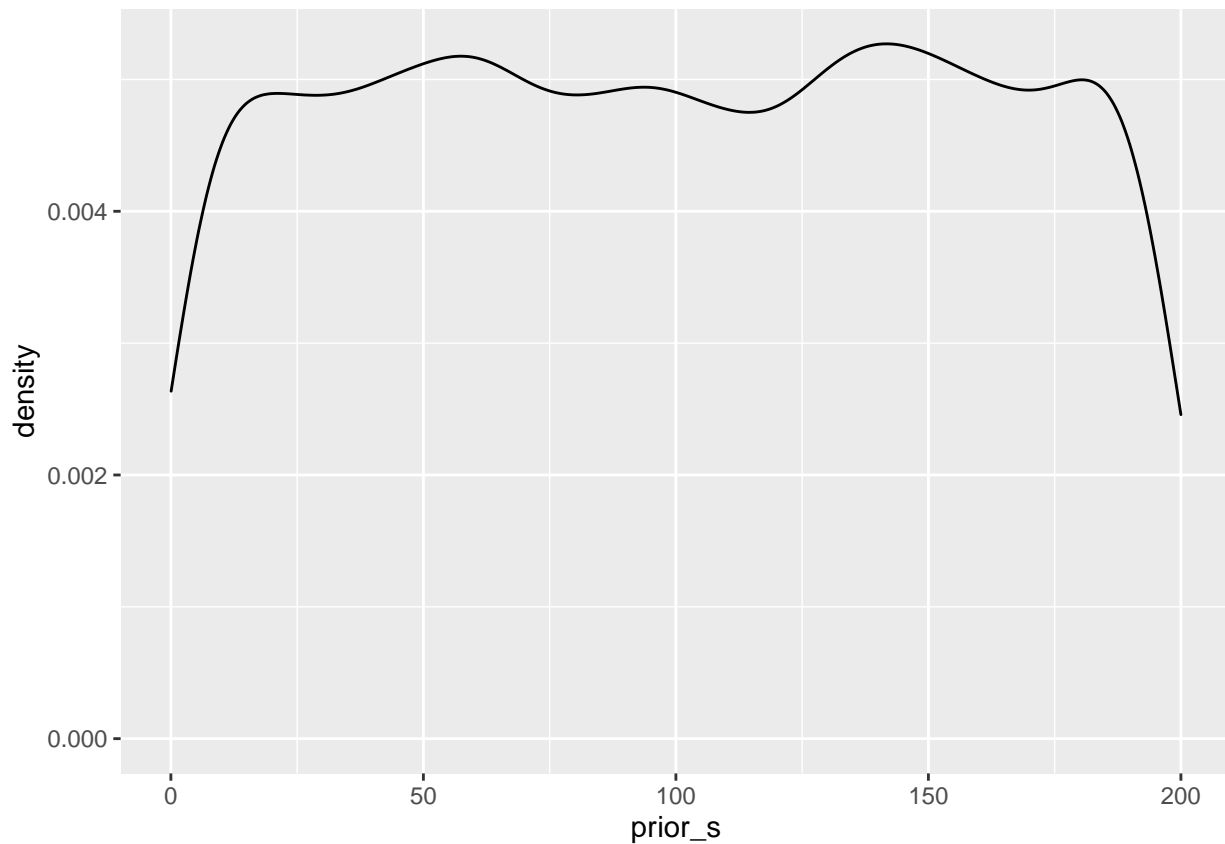
# Take 10000 samples from the s prior
prior_s <- runif(n=10000, min=0, max=200)

# Store samples in a data frame
samples <- data.frame(prior_m, prior_s)

# Density plots of the prior_m & prior_s samples
ggplot(samples, aes(x = prior_m)) +
  geom_density()
```



```
ggplot(samples, aes(x = prior_s)) +  
  geom_density()
```



The distributions of these random samples approximate the features of your Normal prior for m and Uniform prior for s .

Sleep study data

Researchers enrolled 18 subjects in a sleep deprivation study. Their observed `sleep_study` data are loaded in the workspace. These data contain the `day_0` reaction times and `day_3` reaction times after 3 sleep deprived nights for each subject.

You will define and explore `diff_3`, the observed difference in reaction times for each subject. This will require the `mutate()` & `summarize()` functions. For example, the following would add variable `day_0_s`, `day_0` reaction times in seconds, to `sleep_study`:

```
sleep_study <- sleep_study %>%  
  mutate(day_0_s = day_0 * 0.001)
```

You can then `summarize()` the `day_0_s` values, here by their minimum & maximum:

```
sleep_study %>%  
  summarize(min(day_0_s), max(day_0_s))
```

```
sleep_study <- readr::read_csv('data/sleep_study.csv')
```

```
## Parsed with column specification:
## cols(
##   subject = col_double(),
##   day_0 = col_double(),
##   day_3 = col_double()
## )
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

Instructions

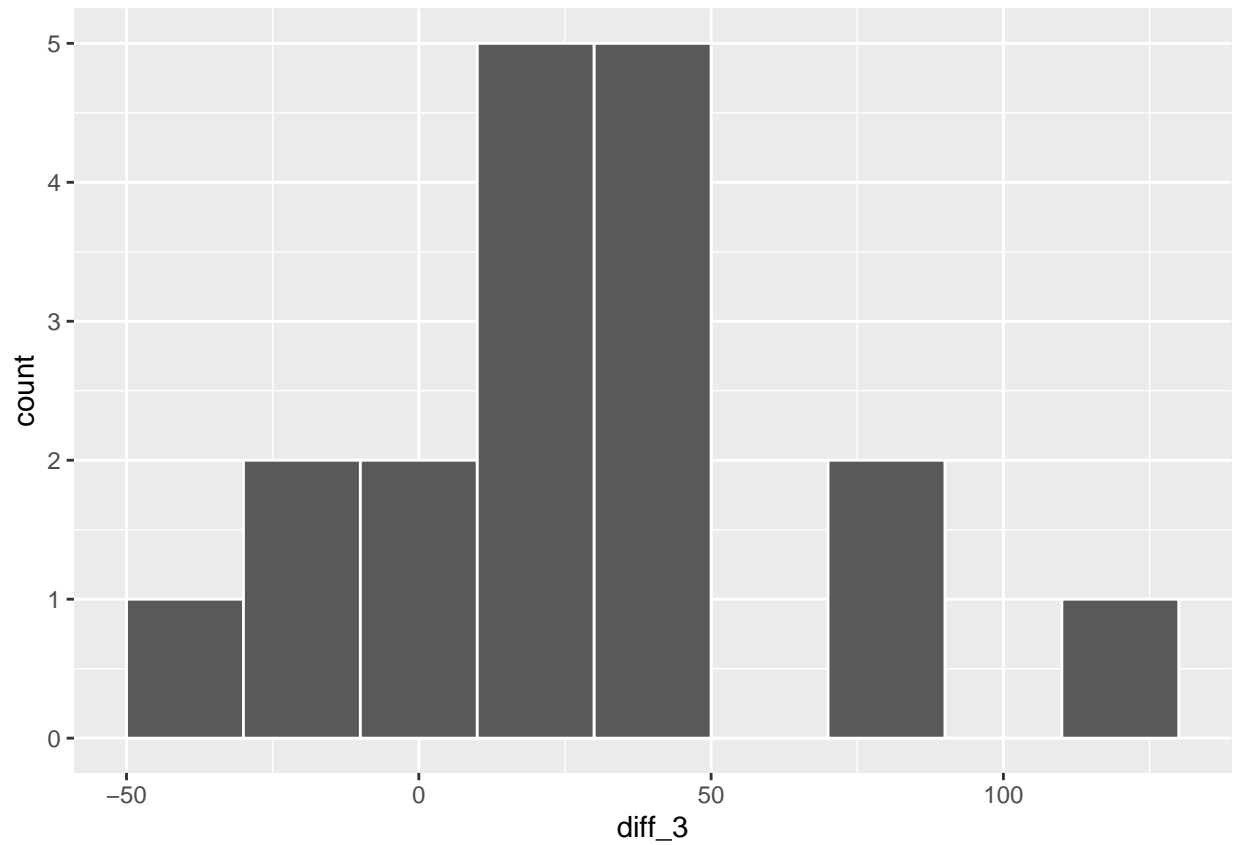
+ Check out the first 6 rows of sleep_study. + Define a new sleep_study variable diff_3, the day_3 minus the day_0 reaction times. + Use ggplot() with a geom_histogram() layer to construct a histogram of the diff_3 data. + summarize() the mean and standard deviation of the diff_3 observations.

```
# Check out the first 6 rows of sleep_study
sleep_study[1:6,]
```

```
## # A tibble: 6 x 3
##   subject day_0 day_3
##   <dbl> <dbl> <dbl>
## 1     308  250.  321.
## 2     309  223.  205.
## 3     310  199.  233.
## 4     330  322.  285.
## 5     331  288.  320.
## 6     332  235.  310.
```

```
# Define diff_3
sleep_study <- sleep_study %>%
  mutate(diff_3 = day_3 - day_0)

# Histogram of diff_3
ggplot(sleep_study, aes(x = diff_3)) +
  geom_histogram(binwidth = 20, color = "white")
```



```
# Mean and standard deviation of diff_3
sleep_study %>%
  summarize(mean(diff_3), sd(diff_3))
```

```
## # A tibble: 1 x 2
##   'mean(diff_3)' 'sd(diff_3)'
##       <dbl>       <dbl>
## 1       26.3       37.2
```