Stat 341 – Homework 7

Gloria Grace

March 24, 2023

8E1

(1) Bread dough rises because of yeast. For this one, the temperature can can lead to an interaction effect between the dough rises and yeast.

(2) Education leads to higher income. A third variable of this can be someone's IQ, when someone has low iq or hard to adapt and absorb education, it might affect their knowledge which could also affect their income.

(3) Gasoline makes a car go. For this one, a third variable can be wheels because technically, a car without wheels cannot go anywhere.

8E2

Out of all the options, the first explanation, is the only one that invokes an interaction because in order to get caramelization, it needed low heat to make sure it doesn't dried out.

Whereas in the second statement, it is stating that either cylinders or a better fuel injector will result in faster speed.

Same with number three, which is presented with an "or." People acquire beliefs from their parents or their friends. Which deliver the message that both don't have interaction with each other (parents or friends).

For the last one, it included the word or again which means that there is no indication that the influence of one factor impacts the influence of the other.

N1

data

```
library(rethinking)
data("Wines2012")
Wines2012 <- Wines2012 |>
  # make a standardized version of the score variable
  mutate(score.std = as.numeric(scale(score)),
         # categorical versions of some variables originally coded 0/1
         wine.origin = ifelse(wine.amer == 1, 'USA', 'Other'),
         judge.nationality = ifelse(judge.amer == 1, 'USA', 'France'),
         # numeric index versions of the same categorical variables
         origin.ix = as.numeric(factor(wine.origin)),
         judge.ix = as.numeric(factor(judge.nationality))
         )
```

```
Wines2012 <- Wines2012 |>
  mutate(Origin_Judge = interaction(wine.origin, judge.nationality),
         Origin_Judge_ix = as.numeric(Origin_Judge)) |>
```

```
  drop_na(Origin_Judge, score.std)
glimpse(Wines2012)
```

```
## Rows: 180
## Columns: 13
## $ judge             <fct> Jean-M Cardebat, Jean-M Cardebat, Jean-M Cardebat, J~
## $ flight            <fct> white, white, white, white, white, white, white, whi~
## $ wine              <fct> A1, B1, C1, D1, E1, F1, G1, H1, I1, J1, A1, B1, C1, ~
## $ score             <dbl> 10.0, 13.0, 14.0, 15.0, 8.0, 13.0, 15.0, 11.0, 9.0, ~
## $ wine.amer         <int> 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0~
## $ judge.amer        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1~
## $ score.std         <dbl> -1.57660412, -0.45045832, -0.07507639, 0.30030555, -~
## $ wine.origin       <chr> "USA", "USA", "Other", "Other", "USA", "USA", "USA",~
## $ judge.nationality <chr> "France", "France", "France", "France", "France", "F~
## $ origin.ix         <dbl> 2, 2, 1, 1, 2, 2, 2, 1, 2, 1, 2, 2, 1, 1, 2, 2, 2, 1~
## $ judge.ix          <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2~
## $ Origin_Judge      <fct> USA.France, USA.France, Other.France, Other.France, ~
## $ Origin_Judge_ix   <dbl> 2, 2, 1, 1, 2, 2, 2, 1, 2, 1, 4, 4, 3, 3, 4, 4, 4, 3~
```

Description

$$\text{score.std}_i \sim \text{Normal}(\mu, \sigma)$$

For some reason this won't work (rmd won't knit)even though I put dollar signs on them. ## mu_i = beta_1 text{[Origin_Judge_ix]}

$$\mu_i = \beta_1$$

$$\beta_1 \sim \text{Normal}(0.1, 0.25)$$

$$\sigma \sim \text{Normal}(\text{mean}_2 = 0, \text{sd}_2 = 1)$$

Defining the interaction

The interactions of the variable `wine.origin` and `judge.nationality` means that the effect of `wine.origin` variable on the response variable is different at different values of the `judge.nationality` variable. Adding a term to the model in which the two predictor variables are multiplied tests this.

Priors

Rationale:

My rational is that I pick a number after looking at the levels of the variable which is there is 4 options.

Fit

```
levels(Wines2012$Origin_Judge)
```

```
## [1] "Other.France" "USA.France"   "Other.USA"    "USA.USA"
```

```
wine_model_list <- alist(
  score.std ~dnorm(mu, sigma),
  mu <- b1[Origin_Judge_ix],
  b1[Origin_Judge_ix] ~ dnorm(0, 0.25),
  sigma ~ dlnorm(0,1)
)
```

```
wine_int_model <- quap(
  flist = wine_model_list,
  data = Wines2012
)
```

```
wine_ps <- extract.samples(wine_int_model) |>
  data.frame()
glimpse(wine_ps)
```

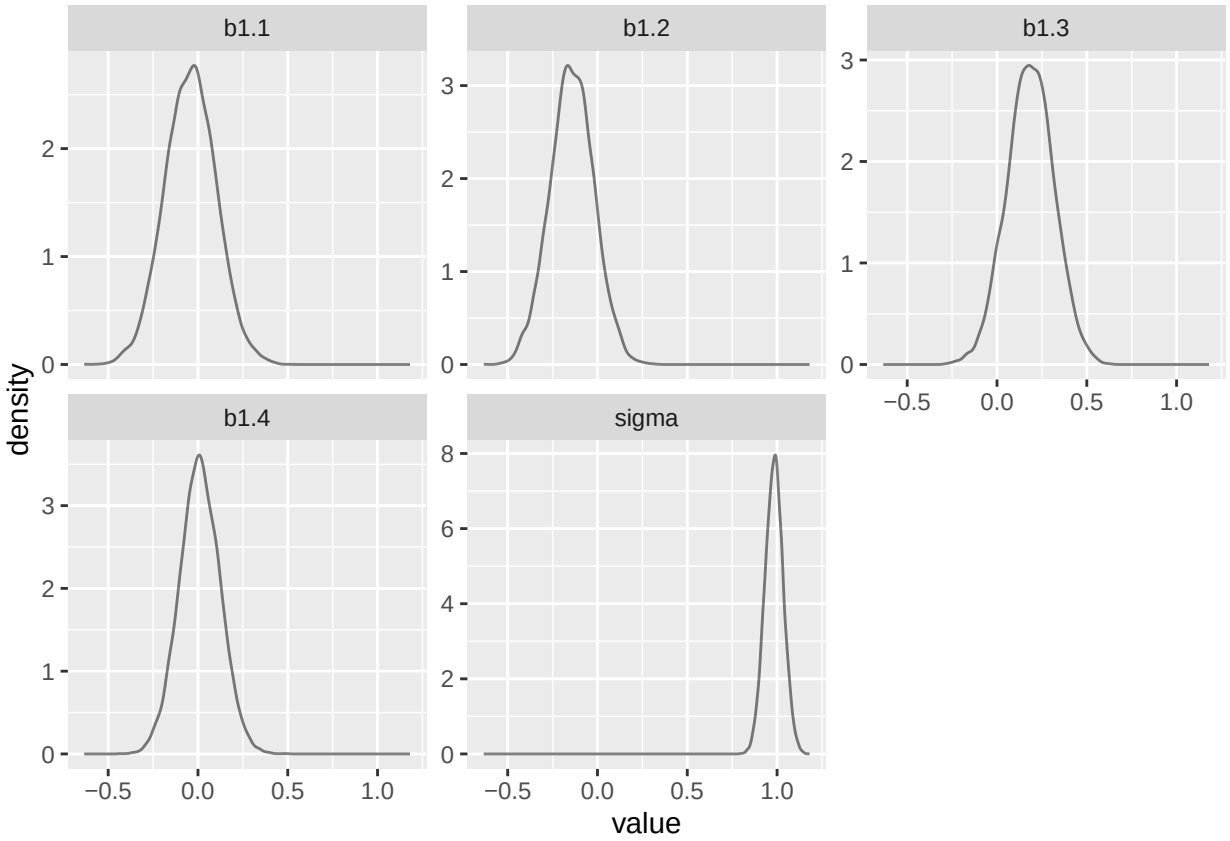
```
## Rows: 10,000
## Columns: 5
## $ sigma <dbl> 1.0206370, 1.0217999, 1.0145298, 0.9085103, 1.0767725, 0.9460176~
## $ b1.1  <dbl> -0.18703307, -0.06333042, -0.01885087, -0.13928547, -0.04641589,~
## $ b1.2  <dbl> -0.40681804, -0.15016277, -0.27487883, -0.13290013, 0.05113117, ~
## $ b1.3  <dbl> 0.12182365, 0.06315970, 0.31829325, 0.19403963, 0.18742070, 0.10~
## $ b1.4  <dbl> -0.1324339794, -0.0711874515, -0.0869608184, 0.0423348217, 0.081~
```

```
wine_ps_plot <- pivot_longer(wine_ps, cols = everything())
glimpse(wine_ps_plot)
```

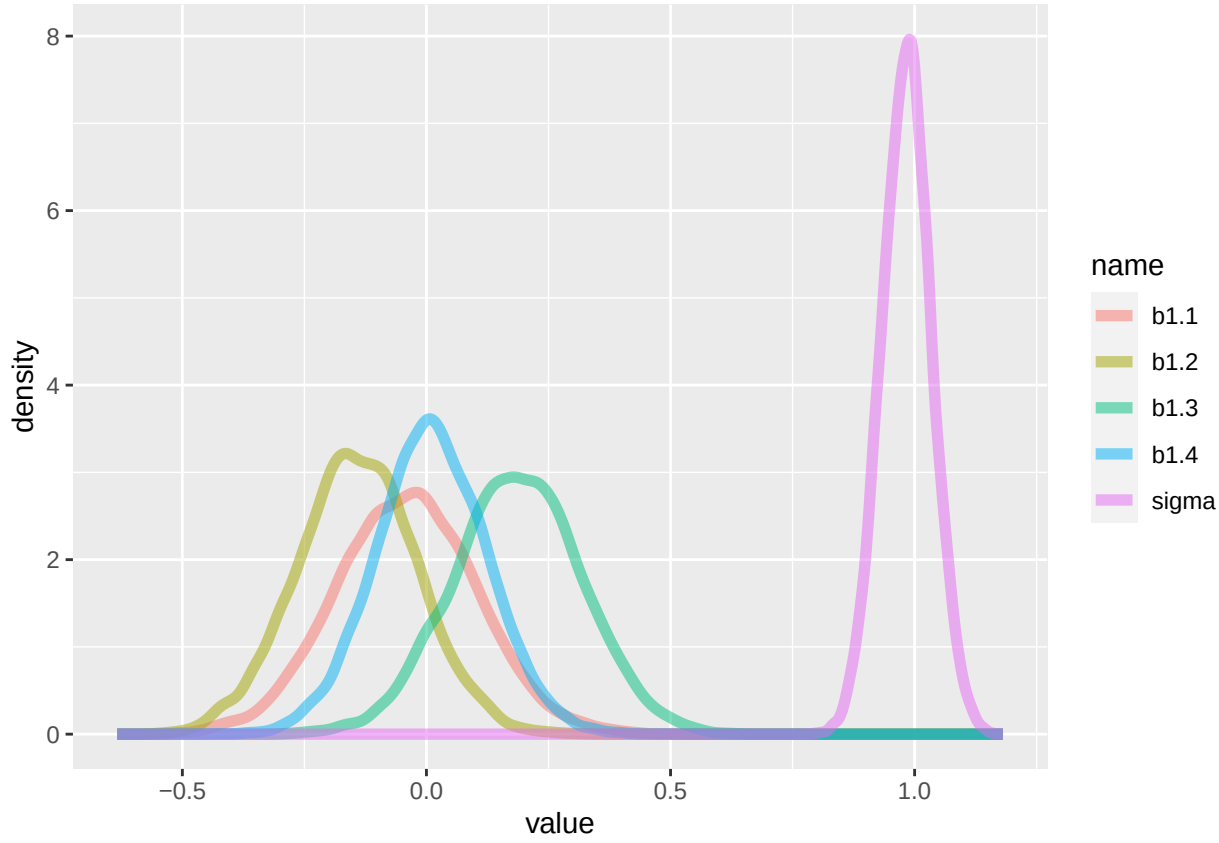
```
## Rows: 50,000
## Columns: 2
## $ name  <chr> "sigma", "b1.1", "b1.2", "b1.3", "b1.4", "sigma", "b1.1", "b1.2"~
## $ value <dbl> 1.02063704, -0.18703307, -0.40681804, 0.12182365, -0.13243398, 1~
```

```
gf_dens(~value, data = wine_ps_plot) |>
  gf_facet_wrap(~name, scales = "free_y")
```

```
## Warning: `stat(density)` was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(density)` instead.
```




```
gf_dens(~value, color = ~name, data = wine_ps_plot, size = 2)
```




```
levels(Wines2012$Origin_Judge)
```

```
## [1] "Other.France" "USA.France"   "Other.USA"    "USA.USA"
```

Interpret

Based on the plot, when the origin is USA and the judge is from france, it has the lowest value for the score.std whereas the highest value goes to the origin other than USA and USA judge.