

Modeling with uncertainty: A Bayesian parameter estimation tutorial

Introduction

Technical objective

Learn the basics of Bayesian parameter estimation, a Bayesian data analysis technique, and how it differs from frequentist parameter estimation

Substantive research question

How do emotion regulation strategies moderate the association between sleep and negative emotions?

Bayesian parameter estimation

In this tutorial, we will be applying the Bayesian lens to the estimation of parameters in a multilevel linear model. In **Bayesian parameter estimation**, we view parameter estimates as probability distributions, as opposed to point values.

Emotion regulation

In the following Bayesian data analysis models, we will examine two emotion regulation strategies, cognitive reappraisal and expressive suppression, and whether they moderate the association between hours of sleep at night and negative affect the following day.

Briefly, in *cognitive reappraisal*, one changes the way in which they think about an emotionally-inducing event in order to alter their emotional response to that event. For example, instead of thinking about an exam as a difficult and consequential test of their abilities, a student might choose to frame the exam as an opportunity to show what they have learned. In *expressive suppression*, one regulates their emotions by inhibiting their external displays of emotion. For example, a student who feels nervous about an upcoming exam may stop themselves from showing a worried look on their face and telling their friends that they feel nervous. More information on these two emotion regulation strategies can be found in Gross & John (2003).

The present study: Emotion regulation, sleep, and negative affect

In our analysis, we will be looking at the relationships among these emotion regulation strategies, sleep, and negative affect, also known as negative emotion. An individual's emotion regulation and coping strategies are thought to moderate the impact of stress on sleep quality (Kahn et al., 2013). We might also be interested in whether emotion regulation strategies moderate the impact of sleep on negative emotions the following day. In other words, does a particular emotion regulation strategy help us manage our negative emotions after a night of poor sleep?

While we may not be able to precisely pin down the causal structure of the interactions among emotion regulation, sleep, and negative affect with our current data, we can run a regression analysis to get some preliminary insights. Our data come from the AMIB study, a multiple time-scale study of college students (Ram et al., 2012). In particular, we will be looking at students' daily self-reports of their sleep and negative affect over the course of eight days, as well as their dispositional reappraisal and suppression scores. The data can be downloaded from <https://thechangelab.stanford.edu/collaborations/the-amib-data/>.

Preliminaries

Load libraries and data

```
library(brms)
library(loo)
library(psych)
library(tinytex)
library(tidyverse)
```

We read in the data from the online repository.

```
# set filepath for data file
filepath <-
  "https://raw.githubusercontent.com/LRI-2/Data/main/ILD/AMIBshare_persons_2019_0501.csv"
# read in the .csv file using the url() function
AMIB_persons <- read.csv(file=url(filepath),header=TRUE)
```

```
# set filepath for data file
filepath <-
  "https://raw.githubusercontent.com/LRI-2/Data/main/ILD/AMIBshare_daily_2019_0501.csv"
# read in the .csv file using the url() function
AMIB_daily <- read.csv(file=url(filepath),header=TRUE)
```

Data manipulation

We merge the daily-level variables (participant ID, day, hours of sleep, and negative affect) and the person-level variables (participant ID, reappraisal score, suppression score) into a single dataset.

```
# subset to day-level variables of interest
bpe_daily <- AMIB_daily[,c("id", "day", "slphrs", "negaff")]
# subset to person-level variables of interest
bpe_persons <- AMIB_persons[,c("id", "erq_reap", "erq_supp")]
# merge day- and person-level data
bpe_data <- merge(bpe_daily, bpe_persons, by = "id")
```

We center the emotion regulation questionnaire scores to facilitate the interpretation of model results.

```
# center the erq subscale score variables for interpretability
bpe_data$erq_reap_c <- scale(bpe_data$erq_reap, center=TRUE,scale=FALSE)[,1]
bpe_data$erq_supp_c <- scale(bpe_data$erq_supp, center=TRUE,scale=FALSE)[,1]
```

Initial plots

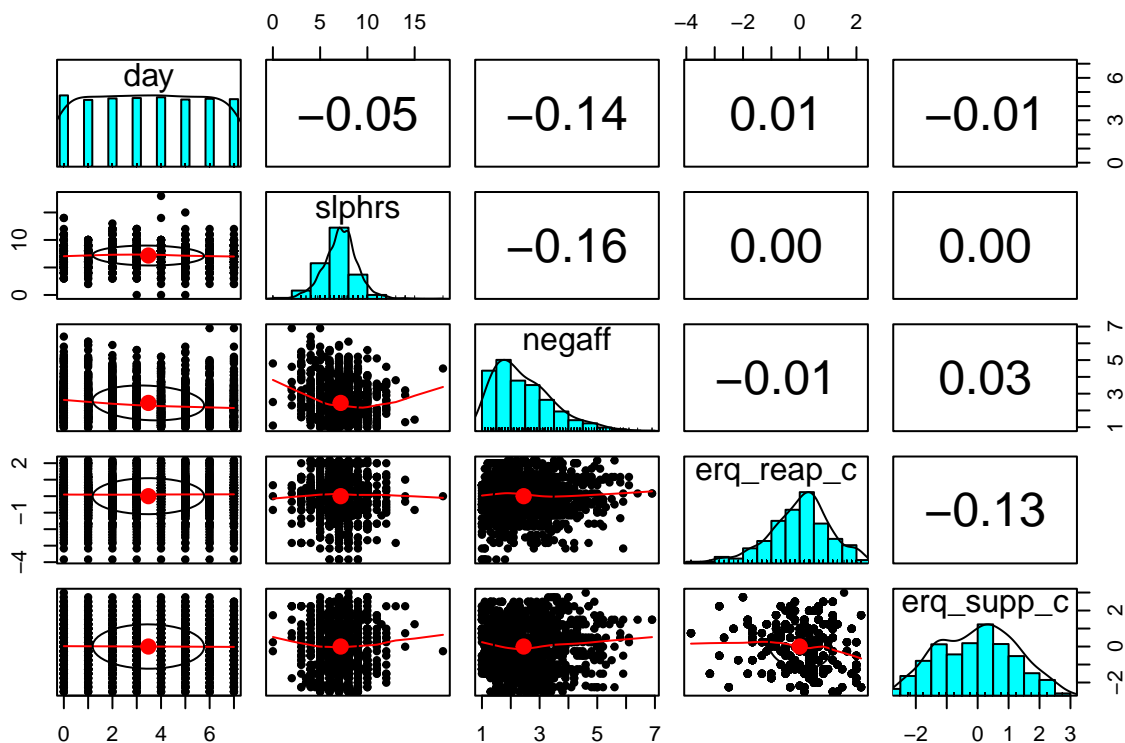
Before running our models, it can be helpful to examine the raw data.

Correlations and distributions

```
# calculate correlations
# dropping the id and non-centered erq columns
cor(bpe_data[,c(-1, -5, -6)], use = "complete.obs")
```

```
##              day      slphrs      negaff      erq_reap_c      erq_supp_c
## day          1.000000000 -0.0480824055 -0.141990043  0.007004748 -0.0210526130
## slphrs       -0.048082405  1.0000000000 -0.155243242  0.001261147 -0.0004183473
## negaff       -0.141990043 -0.1552432419  1.000000000 -0.002317911  0.0246226777
## erq_reap_c    0.007004748  0.0012611469 -0.002317911  1.000000000 -0.1337777409
## erq_supp_c   -0.021052613 -0.0004183473  0.024622678 -0.133777741  1.0000000000
```

```
# plot correlations
pairs.panels(bpe_data[,c(-1, -5, -6)])
```



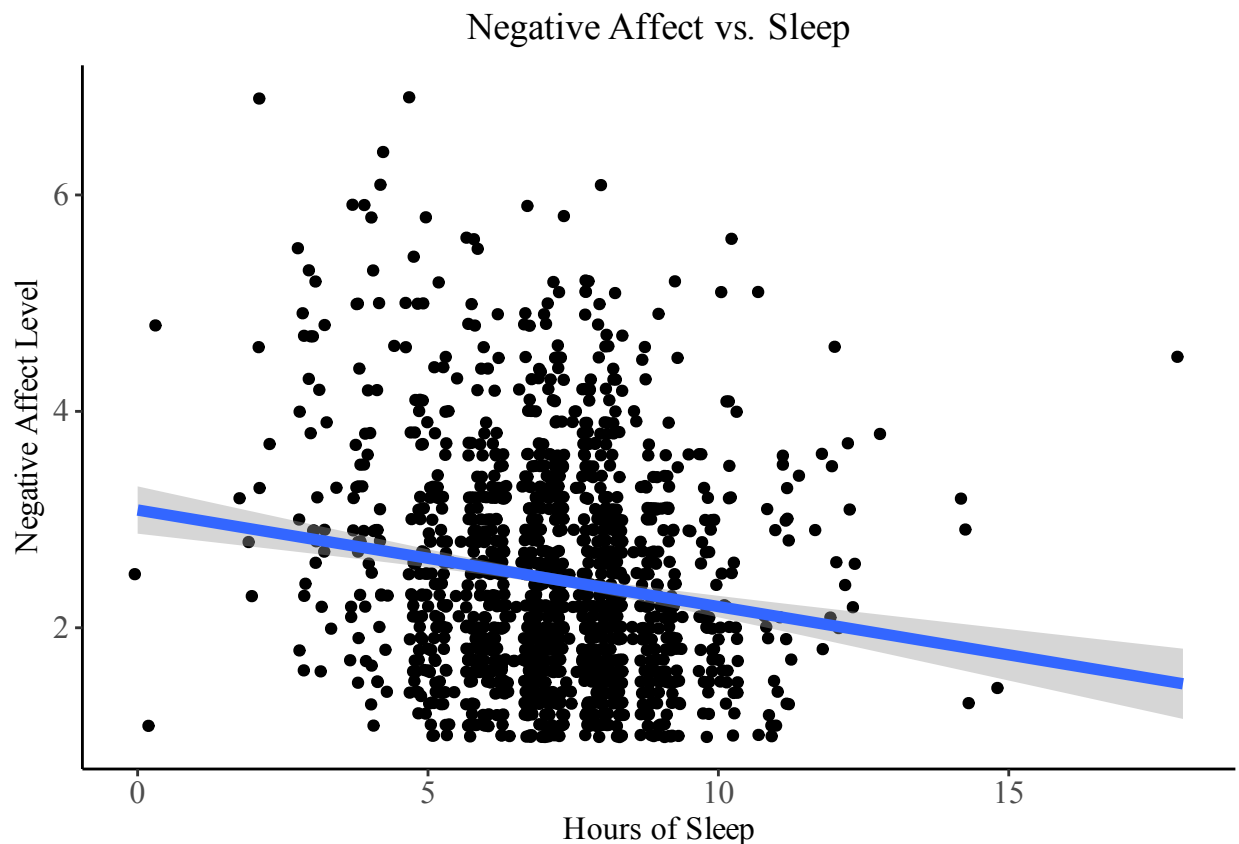
We see relatively weak correlations overall, but note a small negative correlation between hours of sleep and level of negative affect (-0.16), as well as a small negative correlation between cognitive reappraisal and expressive suppression (-0.13). We also observe a small negative correlation between day in study and level of negative affect (-0.14).

The distributions of the variables can be seen on the diagonal. Of particular interest, we note that the distribution of cognitive reappraisal scores has a negative skew, while the distribution of expressive suppression scores is relatively symmetrical.

Linear regression: Negative affect vs. hours of sleep

We will now plot a simple, non-Bayesian linear regression of negative affect vs. hours of sleep to get a better sense of the association between our main predictor variable and our outcome variable.

```
ggplot(data=bpe_data, aes(x=slphrs, y=negaff)) +  
  geom_jitter(width = 0.35) +  
  geom_smooth(method=lm, lty=1, size=2) +  
  xlab("Hours of Sleep") + ylab("Negative Affect Level") +  
  theme_classic() +  
  theme(axis.title=element_text(size=12),  
        axis.text=element_text(size=12),  
        plot.title=element_text(size=14, hjust=.5)) +  
  theme(text = element_text(family = "Times New Roman")) +  
  ggtitle("Negative Affect vs. Sleep")
```



We see that in general, more sleep is associated with lower negative affect. However, there is a wide spread of negative affect values for many sleep durations, suggesting differences in negative affect across individuals.

The Bayesian approach

Before running our Bayesian models, we will briefly recap Bayes' theorem. Bayes' theorem gives us a mathematical expression for the probability of a hypothesis H conditional on (that is, given existing knowledge of) a body of evidence, E (Joyce, 2003/2021). The expression for Bayes' theorem is

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}.$$

The posterior, $P(H|E)$, refers to the probability of the hypothesis being true after having observed the evidence. The likelihood, $P(E|H)$, refers to the probability of observing the evidence in the case where the hypothesis is already true. The prior, $P(H)$, refers to the probability of the hypothesis being true without having observed any evidence. Finally, the normalizing constant, $P(E)$, refers to the probability of the evidence occurring independent of any hypothesis.

In general, Bayes' theorem operates under the assumption that our refined degree of belief in H can be expressed in terms of (i) our prior, or existing, degree of belief in H and (ii) the contribution of newly observed evidence.

Of interest to parameter estimation, Bayes' theorem allows us to calculate the probability distribution of a particular variable. In Bayesian parameter estimation, we use Bayes' theorem to estimate the posterior distribution of a parameter value conditioned on our data, that is, $P(\theta|D)$. Compared to frequentist parameter estimation, which calculates point estimates for parameter values, Bayesian parameter estimation gives us a new way to quantify how uncertain we are about our parameter estimates through the use of probability distributions.

Cognitive reappraisal model

We will first estimate a model using the cognitive reappraisal score as a potential moderator for the relationship between how much one reports sleeping on a particular night and their reported negative affect level the following day.

Note that we are using the `brms` package (Bürkner, 2017) to fit our Bayesian models. This package allows the user to specify prior distributions for parameters. However, with a relatively simple linear model such as this one, we can instead allow `brms` to automatically select uninformative priors for the model's parameters.

Model equation

$$negaff_{it} = \beta_0 + \beta_1 day_t + \beta_2 sleep_{it} + \beta_3 reappraisal_i + \beta_4 sleep_{it} \times reappraisal_i + u_{it}$$

Run model

```
bpe.reap <- brm(negaff ~ 1 + day + slphrs + erq_reap_c + slphrs:erq_reap_c
               + (1 + slphrs + erq_reap_c | id),
               data = bpe_data, family = gaussian(),
               iter = 2000, chains = 4, cores = 4)
summary(bpe.reap)
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: negaff ~ 1 + day + slphrs + erq_reap_c + slphrs:erq_reap_c + (1 + slphrs + erq_reap_c | id)
```

```

## Data: bpe_data (Number of observations: 1421)
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
## total post-warmup draws = 4000
##
## Group-Level Effects:
## ~id (Number of levels: 190)
##
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)      0.97      0.12    0.74    1.22 1.00      996
## sd(slphrs)          0.09      0.02    0.05    0.14 1.01      430
## sd(erq_reap_c)      0.09      0.07    0.00    0.25 1.00      668
## cor(Intercept,slphrs) -0.75      0.09   -0.87   -0.52 1.00     1252
## cor(Intercept,erq_reap_c) 0.04      0.47   -0.84    0.84 1.00     1925
## cor(slphrs,erq_reap_c) -0.14      0.50   -0.92    0.83 1.00     1649
##
##           Tail_ESS
## sd(Intercept)    1213
## sd(slphrs)        534
## sd(erq_reap_c)    1163
## cor(Intercept,slphrs) 1131
## cor(Intercept,erq_reap_c) 2709
## cor(slphrs,erq_reap_c) 2483
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      3.25      0.14    2.98    3.52 1.00     3970     3512
## day            -0.07      0.01   -0.08   -0.05 1.00     9370     2998
## slphrs         -0.08      0.02   -0.11   -0.05 1.00     4019     3211
## erq_reap_c      0.00      0.12   -0.23    0.23 1.00     3635     3021
## slphrs:erq_reap_c 0.00      0.01   -0.03    0.03 1.00     3747     2704
##
## Family Specific Parameters:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma      0.78      0.02    0.75    0.81 1.00     2801     3105
##
## 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).

```

```

bayes_R2(bpe.reap)

```

```

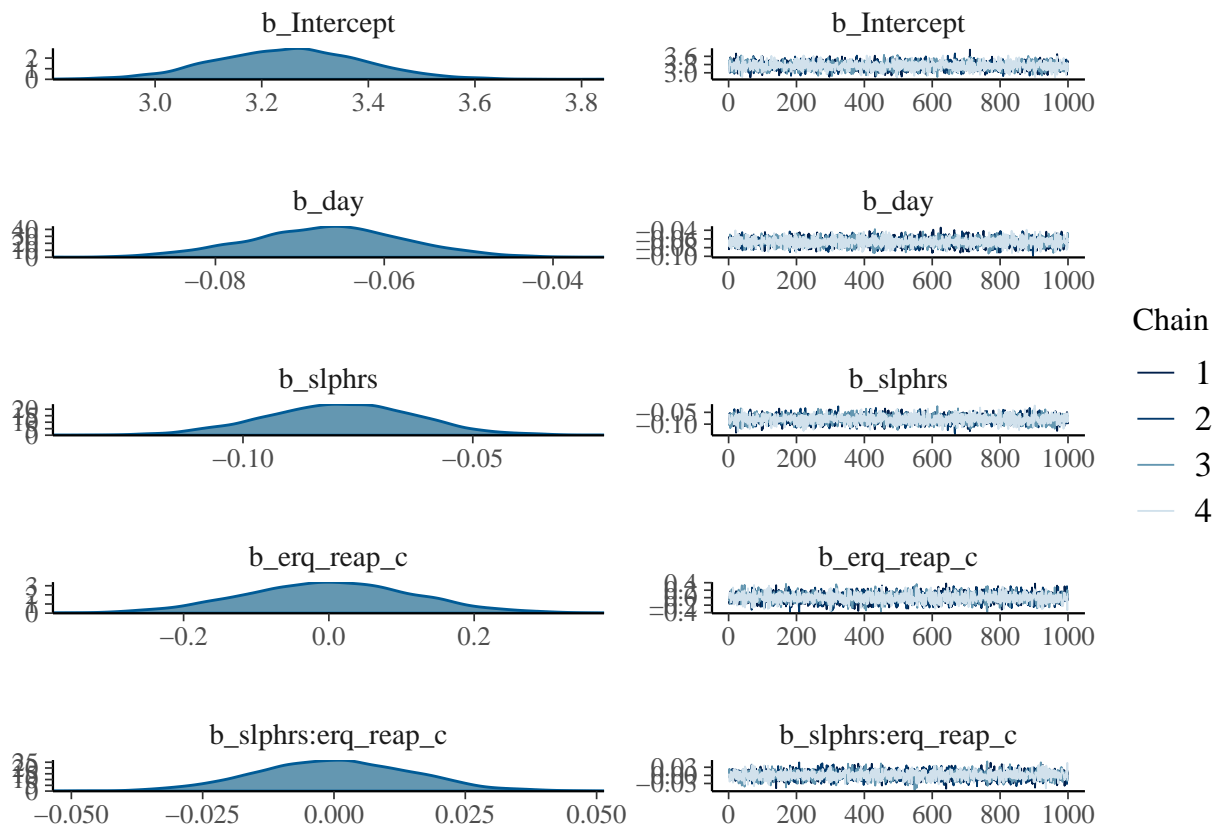
##           Estimate Est.Error      Q2.5      Q97.5
## R2 0.4450501 0.01831485 0.4082889 0.4789184

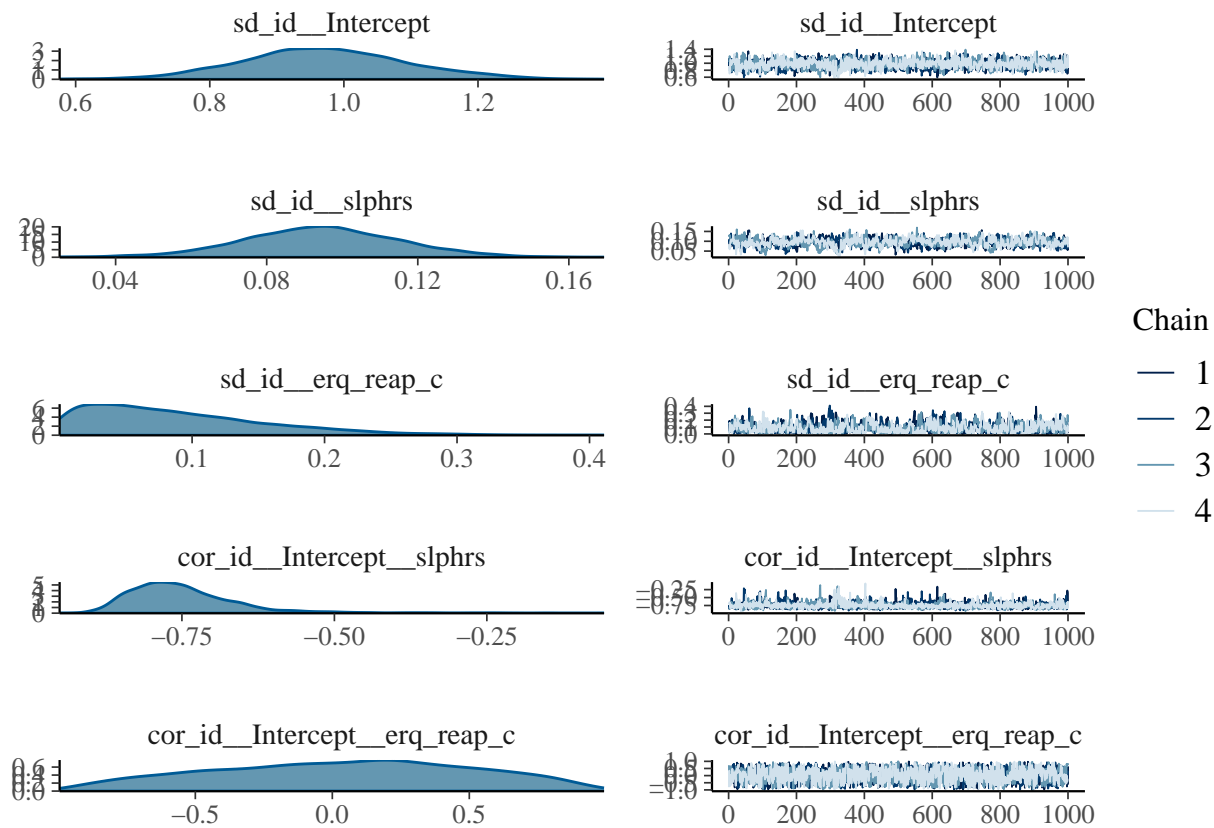
```

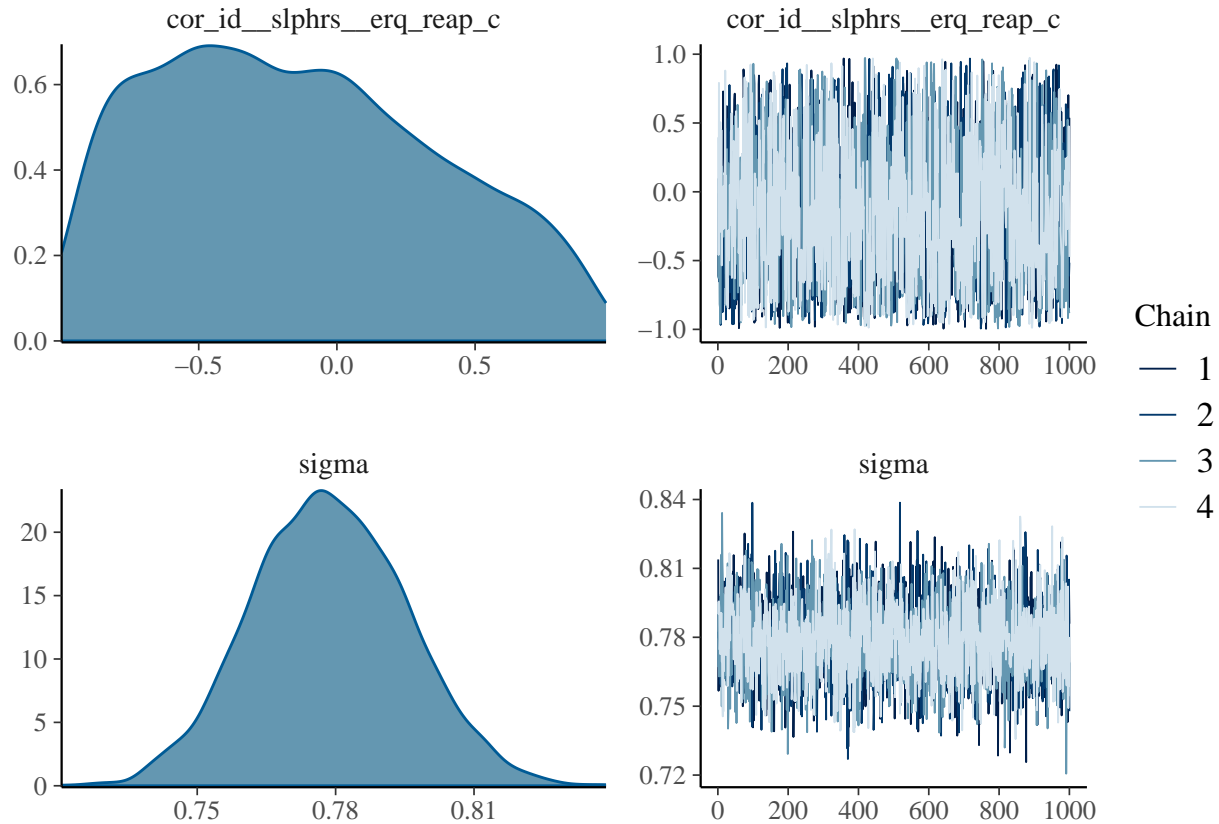
```

plot(bpe.reap)

```







Interpret results

We see that all R-hat values are ≤ 1.01 , which indicates good convergence. In the generated plots, we can view the posterior distributions for our estimated parameters. Since a Bayesian model does not give us p -values like a traditional frequentist model would, it can be tricky to determine the significance of parameter estimates. An informal glimpse at the results suggests that cognitive reappraisal score does not have a significant moderating effect. However, comparing multiple Bayesian models can give us more insight into how to interpret them. We will run a second Bayesian model, then compare our two models.

Visualize model

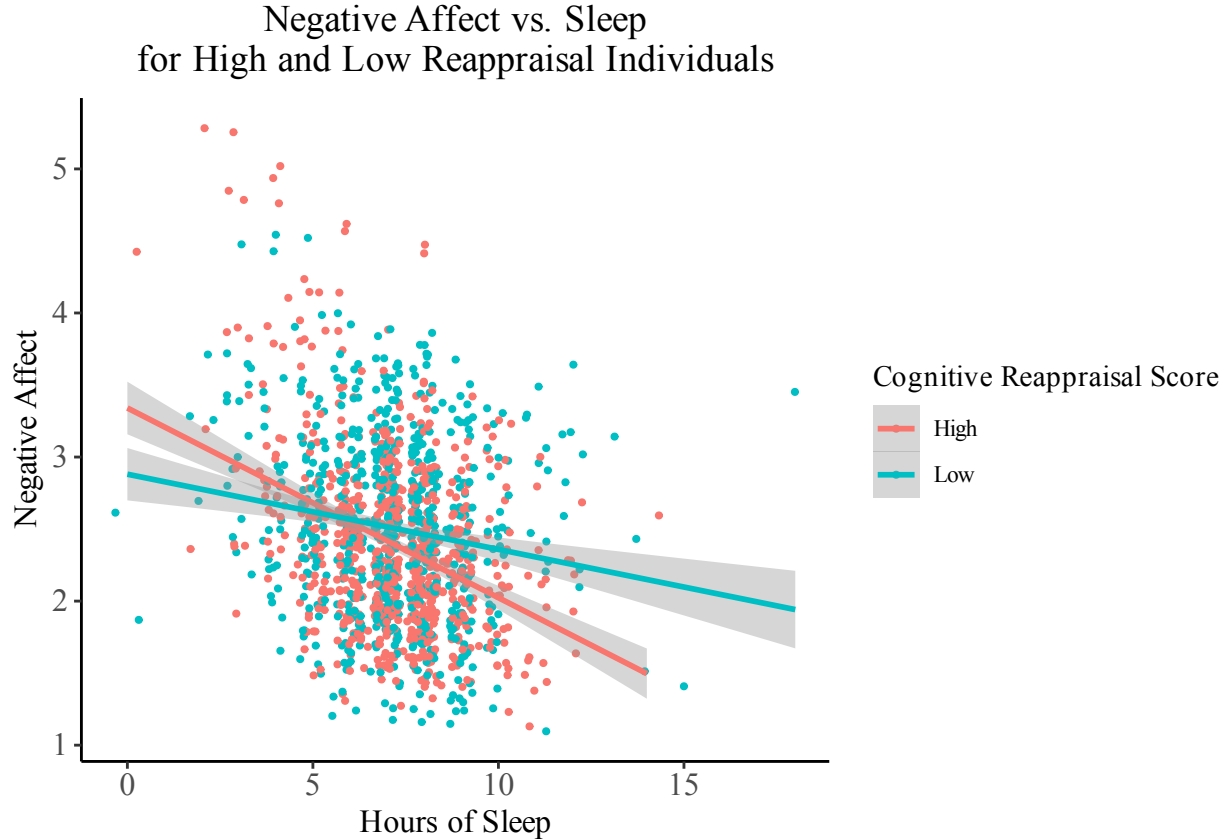
```
# save model predictions to dataset
bpe_data$pred_reap <- predict(bpe.reap, newdata=bpe_data, allow_new_levels=TRUE)
# discretize reappraisal variable for plotting purposes
bpe_data$erq_reap_hl <- ifelse(bpe_data$erq_reap_c >= 0, "High", "Low")

ggplot(data = bpe_data, aes(x = slphrs, y = pred_reap[, "Estimate"],
                             color = erq_reap_hl)) +
  geom_jitter(size = .75, width = 0.35) +
  geom_smooth(method = lm) +
  xlab("Hours of Sleep") +
  ylab("Negative Affect") +
  theme_classic() +
```

```

theme(axis.title=element_text(size=12),
      axis.text=element_text(size=12),
      plot.title=element_text(size=14, hjust=.5)) +
theme(text = element_text(family = "Times New Roman")) +
ggtitle("Negative Affect vs. Sleep\nfor High and Low Reappraisal Individuals") +
labs(colour = "Cognitive Reappraisal Score")

```



We plot the model's predictions for negative affect vs. hours of sleep for “high” (≥ 0) and “low” (< 0) centered cognitive reappraisal scores. We observe that having a higher cognitive reappraisal score may predict a larger difference in negative affect for a fixed difference in hours of sleep. In other words, it appears that those with higher reappraisal scores report negative affect that is more closely related to how much sleep they got the previous night. It also may suggest that cognitive reappraisal is most effective in decreasing negative affect when an individual has gotten sufficient sleep.

Expressive suppression model

Next, we will estimate a very similar model to examine the expressive suppression score as a potential moderator instead of the cognitive reappraisal score.

Model equation

$$negaff_{it} = \beta_0 + \beta_1 day_t + \beta_2 sleephours_{it} + \beta_3 suppression_i + \beta_4 sleephours_{it} \times suppression_i + u_{it}$$

Run model

```
bpe.supp <- brm(negaiff ~ 1 + day + slphrs + erq_supp_c + slphrs:erq_supp_c
               + (1 + slphrs + erq_supp_c|id),
               data = bpe_data, family = gaussian(),
               iter = 2000, chains = 4, cores = 4)
summary(bpe.supp)
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: negaiff ~ 1 + day + slphrs + erq_supp_c + slphrs:erq_supp_c + (1 + slphrs + erq_supp_c | id)
## Data: bpe_data (Number of observations: 1421)
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
## total post-warmup draws = 4000
##
## Group-Level Effects:
## ~id (Number of levels: 190)
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS
sd(Intercept)	0.94	0.13	0.69	1.20	1.00	1361
sd(slphrs)	0.09	0.02	0.05	0.13	1.01	473
sd(erq_supp_c)	0.11	0.07	0.01	0.26	1.00	552
cor(Intercept,slphrs)	-0.73	0.10	-0.87	-0.48	1.00	1768
cor(Intercept,erq_supp_c)	0.34	0.39	-0.53	0.93	1.00	1579
cor(slphrs,erq_supp_c)	-0.15	0.45	-0.90	0.75	1.00	1417

```
##
```

	Tail_ESS
sd(Intercept)	2002
sd(slphrs)	1052
sd(erq_supp_c)	985
cor(Intercept,slphrs)	2019
cor(Intercept,erq_supp_c)	2409
cor(slphrs,erq_supp_c)	2244

```
##
## Population-Level Effects:
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	3.23	0.13	2.97	3.49	1.00	4083	2822
day	-0.07	0.01	-0.08	-0.05	1.00	7999	2807
slphrs	-0.08	0.02	-0.11	-0.05	1.00	3716	3003
erq_supp_c	0.09	0.10	-0.10	0.28	1.00	3600	3316
slphrs:erq_supp_c	-0.01	0.01	-0.03	0.01	1.00	4007	3100

```
##
## Family Specific Parameters:
##
```

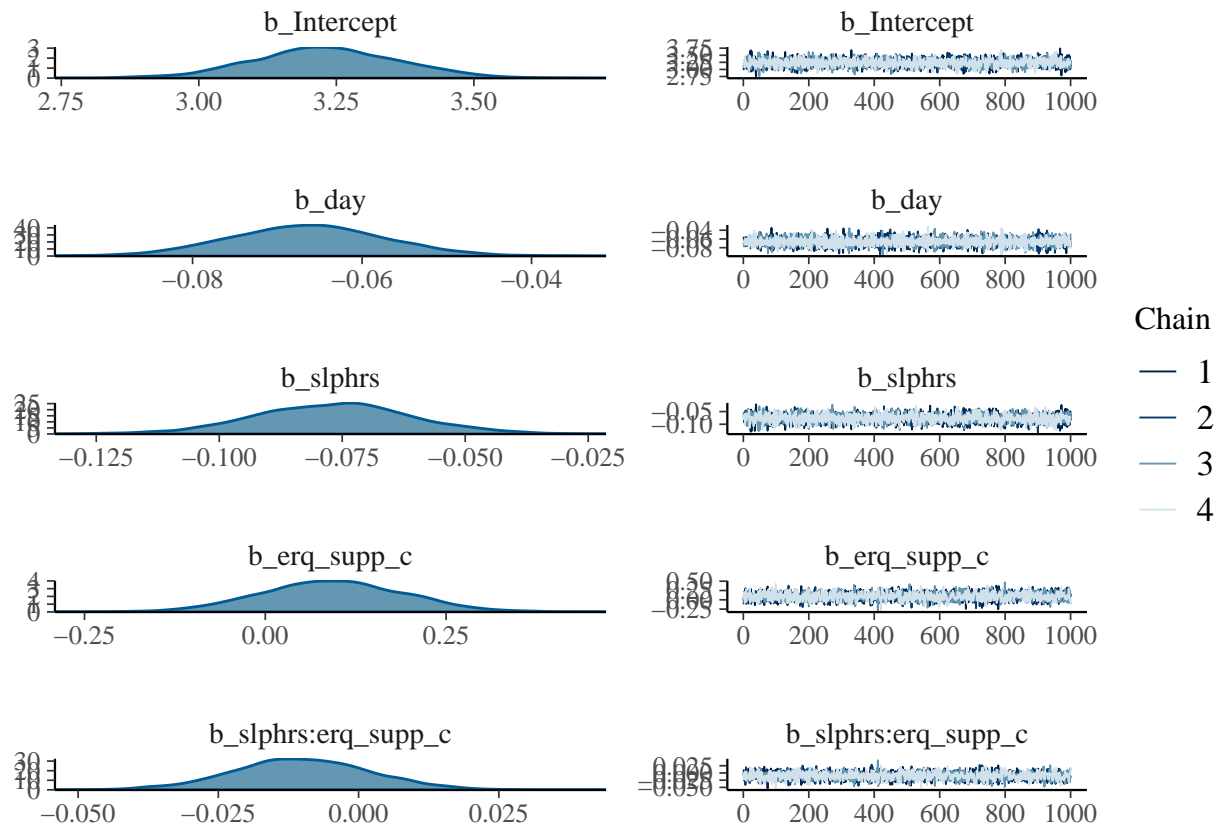
	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	0.78	0.02	0.75	0.81	1.00	3298	2852

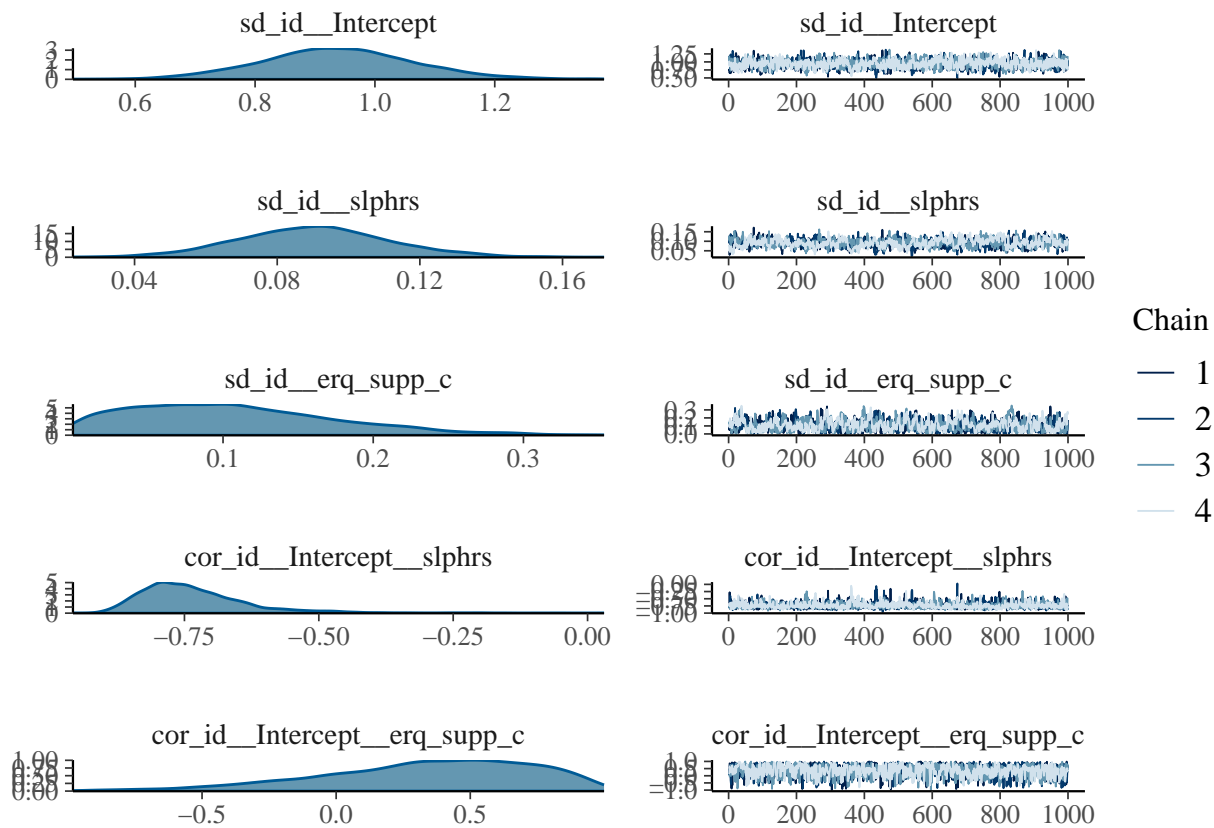
```
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

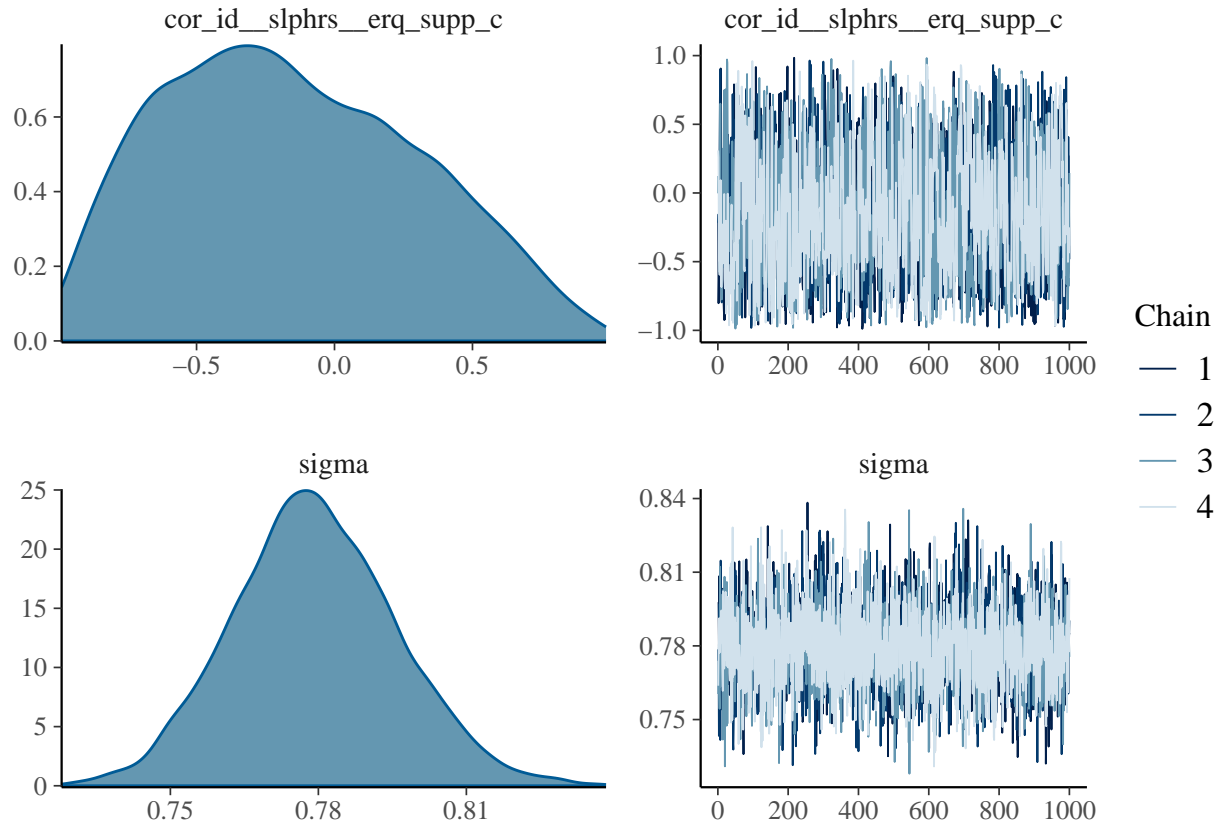
```
bayes_R2(bpe.supp)
```

```
## Estimate Est.Error Q2.5 Q97.5
## R2 0.444175 0.01788962 0.4073789 0.4780995
```

```
plot(bpe.supp)
```







Interpret results

We see that all R-hat values are ≤ 1.01 , which indicates good convergence. In the generated plots, we can view the posterior distributions for our estimated parameters. Since a Bayesian model does not give us p -values like a traditional frequentist model would, it can be tricky to determine the significance of parameter estimates. An informal glimpse at the results suggests that expressive suppression score does not have a significant moderating effect. However, later in the tutorial, we will compare our two models to get a better idea of their comparative fit and significance.

Visualize model

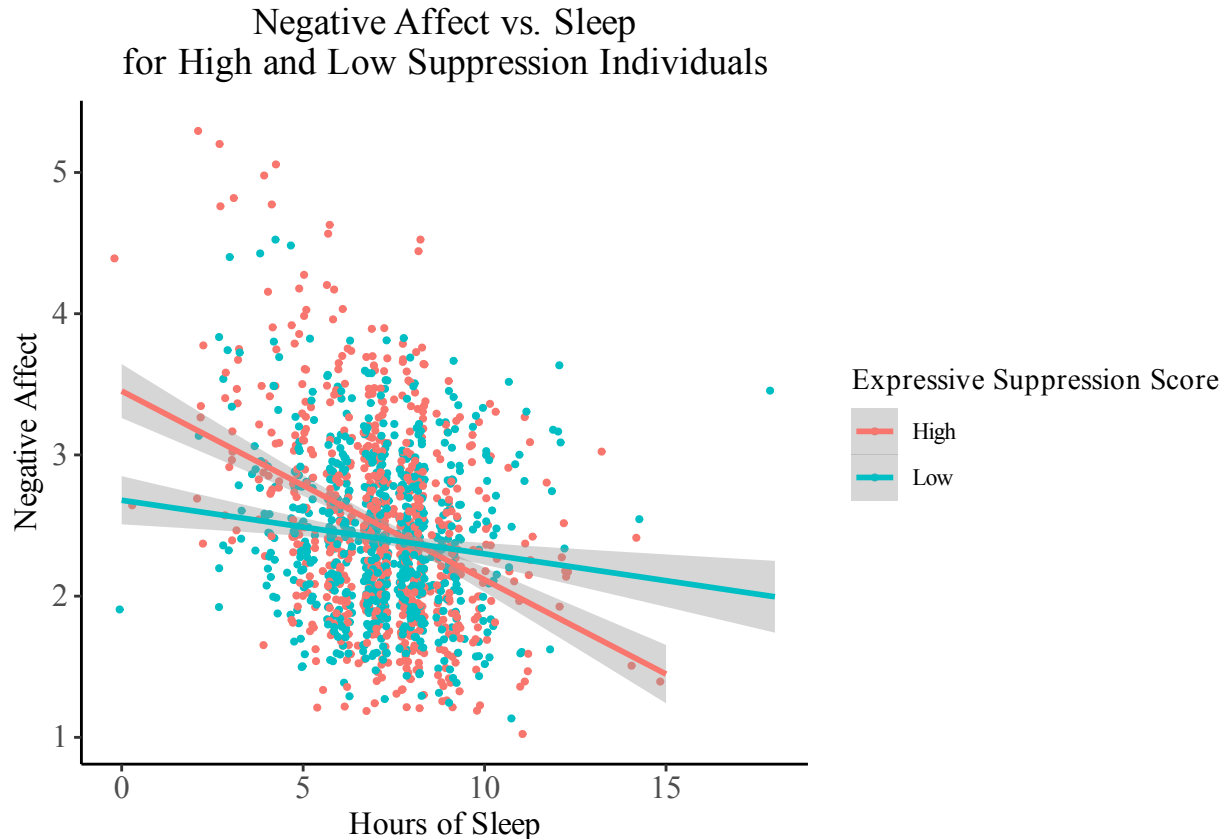
```
# save model predictions to dataset
bpe_data$pred_supp <- predict(bpe.supp, newdata=bpe_data, allow_new_levels=TRUE)
# discretize suppression variable for plotting purposes
bpe_data$erq_supp_hl <- ifelse(bpe_data$erq_supp_c >= 0, "High", "Low")

ggplot(data = bpe_data, aes(x = slphrs, y = pred_supp[, "Estimate"],
                             color = erq_supp_hl)) +
  geom_jitter(size = .75, width = 0.35) +
  geom_smooth(method = lm) +
  xlab("Hours of Sleep") +
  ylab("Negative Affect") +
  theme_classic() +
```

```

theme(axis.title=element_text(size=12),
      axis.text=element_text(size=12),
      plot.title=element_text(size=14, hjust=.5)) +
theme(text = element_text(family = "Times New Roman")) +
ggtitle("Negative Affect vs. Sleep\nfor High and Low Suppression Individuals") +
labs(colour = "Expressive Suppression Score")

```



We plot the model's predictions for negative affect vs. hours of sleep for “high” (≥ 0) and “low” (< 0) centered expressive suppression scores. We observe that having a higher expressive suppression score may predict a larger difference in negative affect for a fixed difference in hours of sleep. In other words, it appears that those with higher suppression scores report negative affect that is more closely related to how much sleep they got the previous night. It also may suggest that expressive suppression, like cognitive reappraisal, is most effective in decreasing negative affect when an individual has gotten sufficient sleep.

We observe that the moderating effects of cognitive reappraisal and expressive suppression appear similar, suggesting that a more general factor is moderating the relationship between hours of sleep and negative affect.

Model comparison

We will now use leave-one-out cross-validation (LOOCV) to compare the relative fits of the cognitive reappraisal model and the expressive suppression model. This type of model comparison is typically done between two different kinds of models using the same data. Our models have the same form but use slightly different data (reappraisal predictor vs. suppression predictor). However, for demonstrative purposes, we will still compare them using LOOCV.

```
bpe.reap <- add_criterion(bpe.reap, c("loo"))
bpe.supp <- add_criterion(bpe.supp, c("loo"))
loo_compare(bpe.reap, bpe.supp)
```

```
##           elpd_diff se_diff
## bpe.supp 0.0         0.0
## bpe.reap 0.0         1.4
```

We observe that the cognitive reappraisal model appears to have a better fit than the expressive suppression model. This indicates that the relationships among variables in the cognitive reappraisal model are better suited to a linear model than those in the expressive suppression model. A LOOCV comparison would likely give us more helpful information if we were comparing, for example, a linear version of the reappraisal model and a nonlinear version of the reappraisal model.

Beyond using formal quantitative techniques like LOOCV for model comparison, it can be helpful to compare the plotted posterior distributions of parameters across the two models. This can help us understand the relative levels of uncertainty we have in those models.

Conclusion

While shifting from a frequentist understanding of parameter estimation to a Bayesian understanding can be challenging, it is important to consider the merits and drawbacks of both approaches. Bayesian parameter estimation can seem unintuitive at first, but it provides us with new tools to describe and analyze the uncertainty inherent in our data. In social science research, being able to incorporate subjectivity into our models can give us a new outlook on phenomena that are often hard to pin down in a standard quantitative framework.

References

- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>
- Joyce, J. (2021). Bayes' theorem (E. N. Zalta, Ed.). *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition). <https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/> (Original work published 2003)
- Kahn, M., Sheppes, G., & Sadeh, A. (2013). Sleep and emotions: Bidirectional links and underlying mechanisms. *International Journal of Psychophysiology*, 89(2), 218–228. <https://doi.org/10.1016/j.ijpsycho.2013.05.010>
- Ram, N., Conroy, D. E., Pincus, A. L., Hyde, A. L., & Molloy, L. E. (2012). Tethering theory to method: Using measures of intraindividual variability to operationalize individuals' dynamic characteristics. In G. Hancock & J. Harring (Eds.), *Advances in longitudinal methods in the social and behavioral sciences* (pp. 81–110). New York: Information Age.