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?

Emotion regulation

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. 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 (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 me manage my negative emotions after a night of poor sleep?

While we may not be able to precisely pin down the causal structure of the interactions between emotion regulation, sleep, and negative affect, 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/.

TODO: add more commentary throughout to guide reader through the steps of the tutorial TODO: use this link for some example plots/analysis of Bayesian MLMs: https://www.rensvandeschoot.com/tutorials/brms-started/

Preliminaries

We begin by loading 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) with 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 data plotting

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

Correlations and distributions

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

```
##                 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)])
```
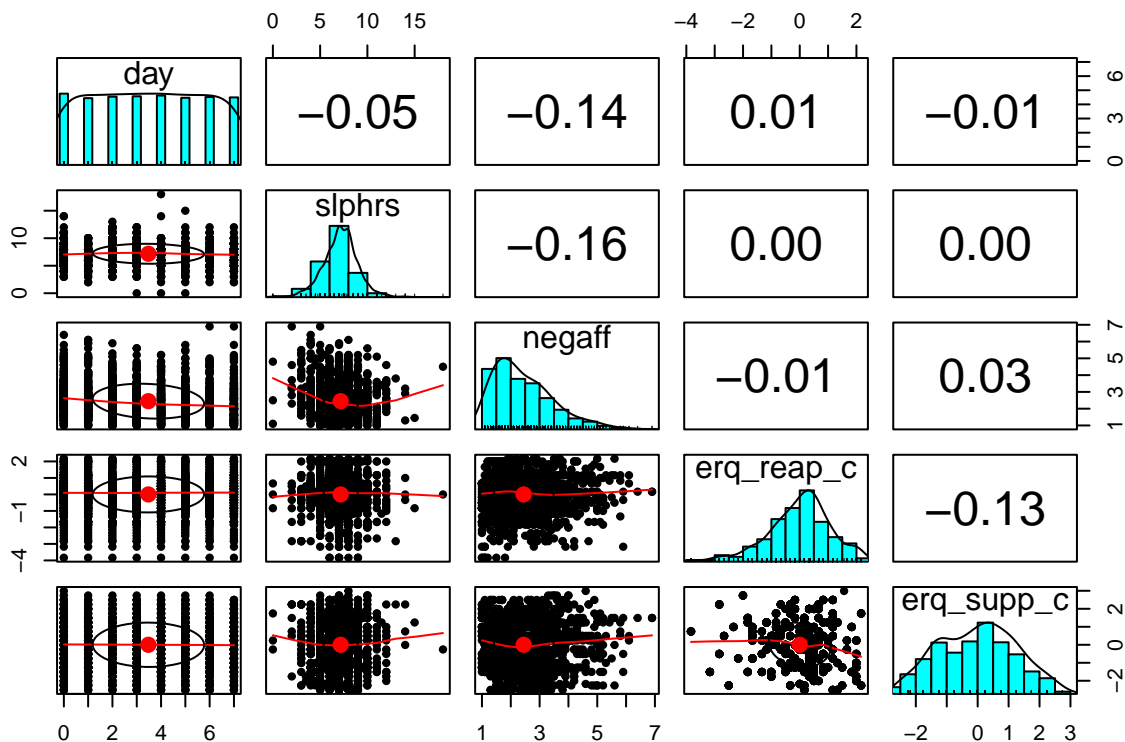

We see relatively weak correlations overall, but note a weak negative correlation between hours of sleep and level of negative affect (-0.16), as well as a weak negative correlation between cognitive reappraisal and expressive suppression (-0.13). We also observe a weak negative correlation between day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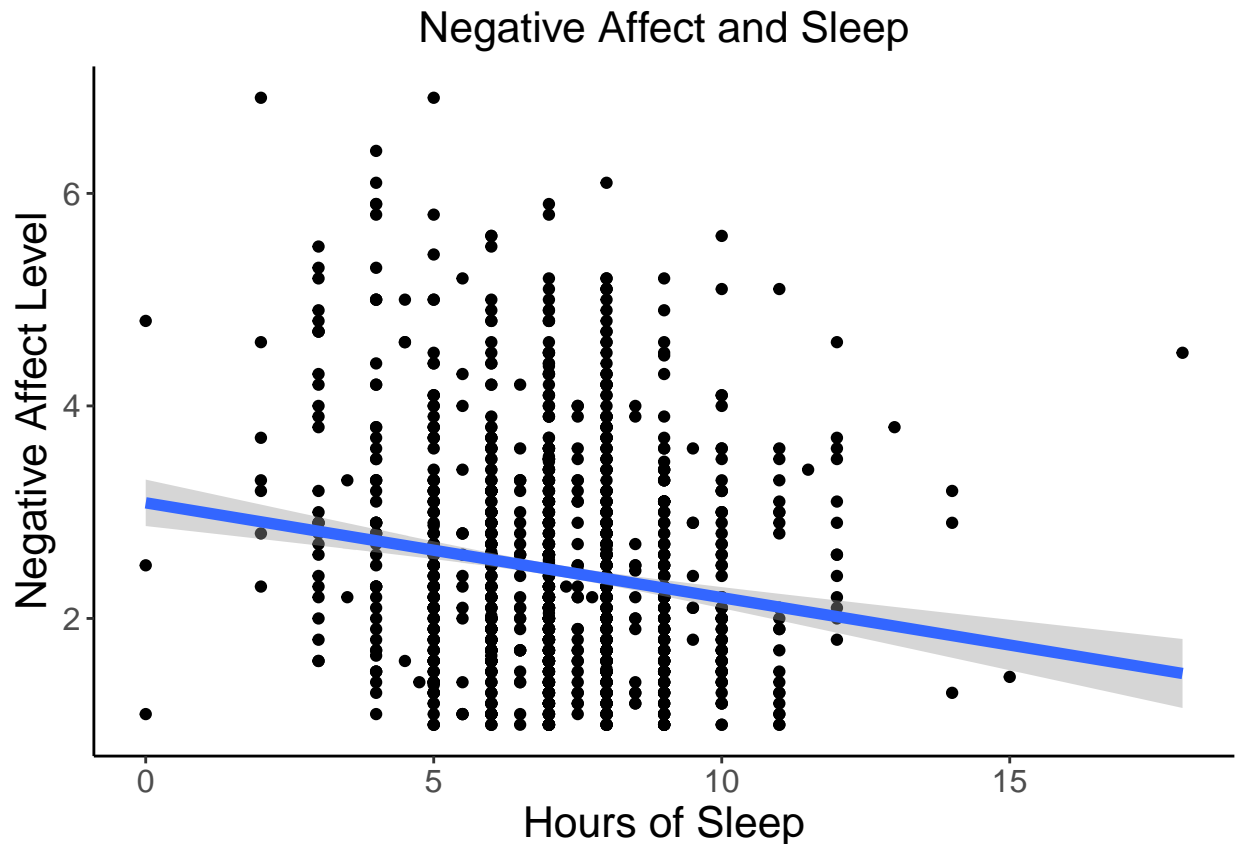
Linear regression: negative affect vs. sleep hours

```
ggplot(data=bpe_data, aes(x=slphrs, y=negaff)) +
  geom_point() +
  geom_smooth(method=lm, lty=1, size=2) +
```

```
xlab("Hours of Sleep") + ylab("Negative Affect Level") +
theme_classic() +
theme(axis.title=element_text(size=16),
      axis.text=element_text(size=12),
      plot.title=element_text(size=16, hjust=.5)) +
ggtitle("Negative Affect and 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.

TODO: add a section going over the basics of Bayesian parameter estimation and how it differs from frequentist parameter estimation; make sure to go into enough mathematical detail

Cognitive reappraisal model

TODO: add model equation in Latex

```
bpe.reap <- brm(negaff ~ 1 + day + slphrs + erq_reap_c + slphrs:erq_reap_c + (1 + slphrs + erq_reap_c|i
  iter = 2000, chains = 4, cores = 4)
```

```
## Warning: Rows containing NAs were excluded from the model.
```

```
## Compiling Stan program...
```

```
## Start sampling
```

```
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 l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)               0.97      0.12     0.74     1.22 1.00     1432
## sd(slphrs)                  0.09      0.02     0.05     0.14 1.01      461
## sd(erq_reap_c)              0.09      0.07     0.00     0.26 1.01      615
## cor(Intercept,slphrs)      -0.75      0.09    -0.87    -0.53 1.01     1402
## cor(Intercept,erq_reap_c)   0.04      0.47    -0.86     0.84 1.00     2047
## cor(slphrs,erq_reap_c)     -0.13      0.51    -0.92     0.83 1.00     1706
##                         Tail_ESS
## sd(Intercept)               2729
## sd(slphrs)                   817
## sd(erq_reap_c)               757
## cor(Intercept,slphrs)       1694
## cor(Intercept,erq_reap_c)   2909
## cor(slphrs,erq_reap_c)      2626
##
## Population-Level Effects:
##                   Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept             3.25      0.14     2.98     3.51 1.00     4585     3689
## day                  -0.07      0.01    -0.09    -0.05 1.00    10040     2558
## slphrs               -0.08      0.02    -0.11    -0.05 1.00     4310     3427
## erq_reap_c           -0.00      0.12    -0.24     0.23 1.00     2796     2911
## slphrs:erq_reap_c     0.00      0.02    -0.03     0.03 1.00     2851     2942
##
## 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     4191     3052
##
## 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).
```
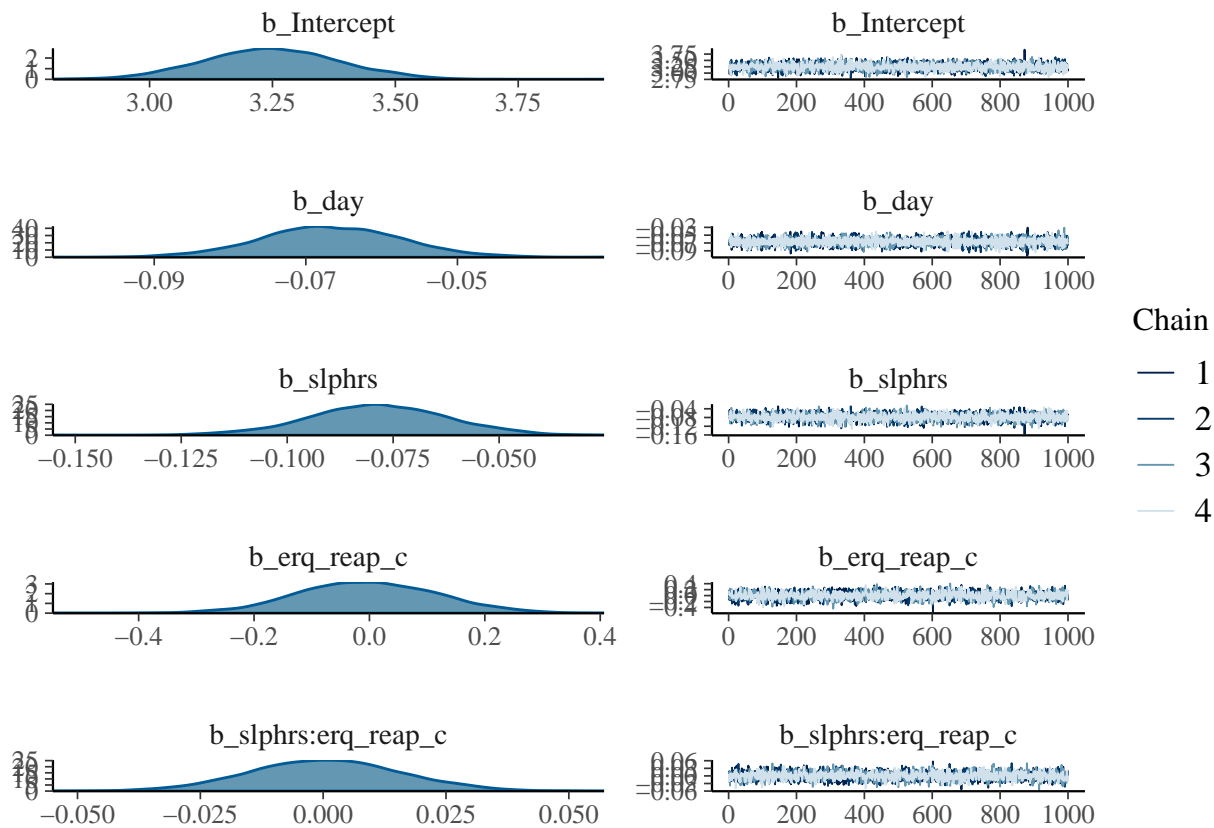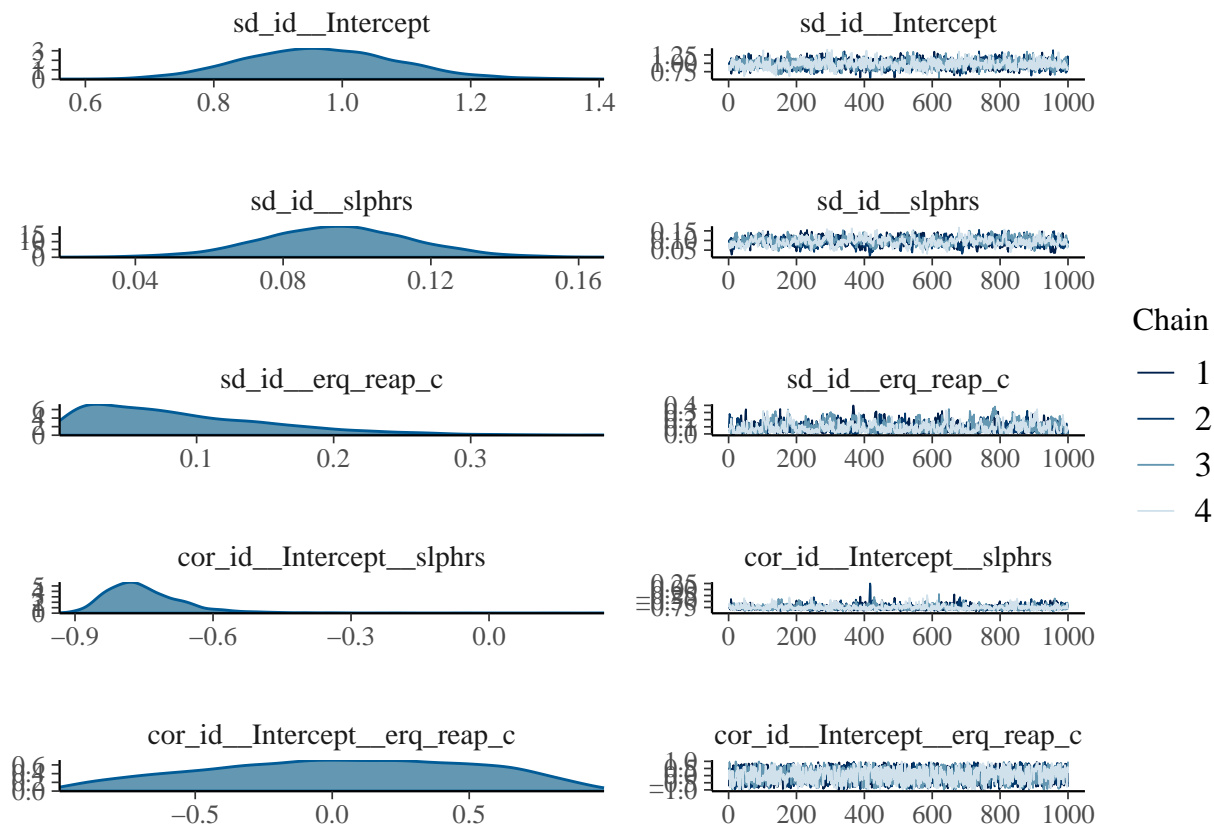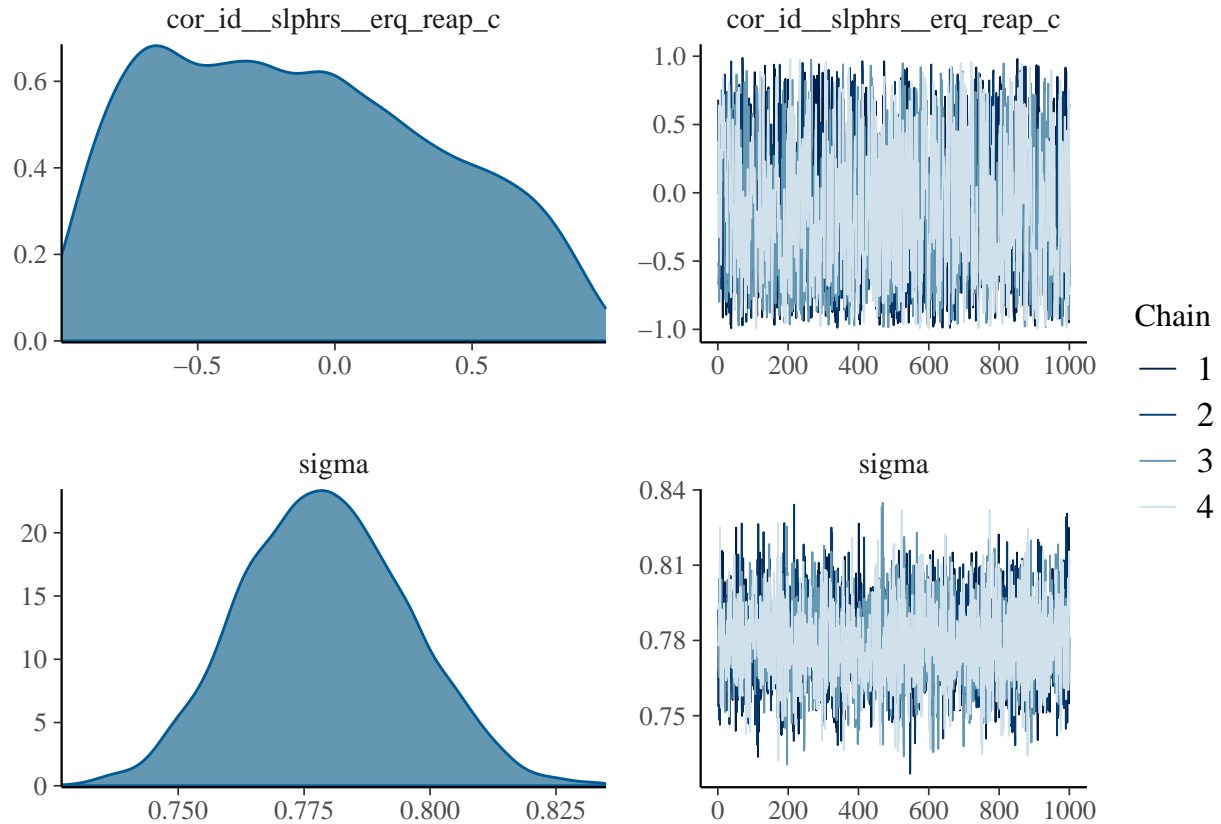
```
bayes_R2(bpe.reap)
```

```
##     Estimate  Est.Error      Q2.5     Q97.5
## R2 0.4441178 0.01838396 0.4060122 0.4788968
```

```
plot(bpe.reap)
```

sd_id__Intercept

sd_id__Intercept

sd_id__slphrs

sd_id__slphrs

Chain

— 1

— 2

— 3

— 4

sd_id__erq_reap_c

sd_id__erq_reap_c

cor_id__Intercept__slphrs

cor_id__Intercept__slphrs

cor_id__Intercept__erq_reap_c

cor_id__Intercept__erq_reap_c

TODO: interpret model results (including Bayesian significance testing) TODO: make a visualization of this model that shows the prototypical individual and all individuals

Expressive suppression model

TODO: add model equation in Latex

```
bpe.supp <- brm(negaff ~ 1 + day + slphrs + erq_supp_c + slphrs:erq_supp_c + (1 + slphrs + erq_supp_c|id
    iter = 2000, chains = 4, cores = 4)
```

```
## Warning: Rows containing NAs were excluded from the model.
```

```
## Compiling Stan program...
```

```
## Start sampling
```

```
## Warning: There were 6 divergent transitions after warmup. See
## https://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
## to find out why this is a problem and how to eliminate them.
```

```
## Warning: Examine the pairs() plot to diagnose sampling problems
```

```
## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quanti
## Running the chains for more iterations may help. See
## https://mc-stan.org/misc/warnings.html#tail-ess
```

```r
summary(bpe.supp)
```

```
## Warning: There were 6 divergent transitions after warmup. Increasing
## adapt_delta above 0.8 may help. See
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: negaff ~ 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     1.19 1.00      991
## sd(slphrs)                0.09      0.02     0.05     0.13 1.01      415
## sd(erq_supp_c)            0.11      0.07     0.01     0.26 1.01      464
## cor(Intercept,slphrs)    -0.73      0.11    -0.87    -0.47 1.00     1346
## cor(Intercept,erq_supp_c) 0.34      0.39    -0.55     0.92 1.00     1680
## cor(slphrs,erq_supp_c)   -0.16      0.45    -0.90     0.75 1.00     1292
##                       Tail_ESS
## sd(Intercept)              946
## sd(slphrs)                 377
## sd(erq_supp_c)             758
## cor(Intercept,slphrs)     1059
## cor(Intercept,erq_supp_c) 2207
## cor(slphrs,erq_supp_c)    1970
##
## Population-Level Effects:
##                   Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept             3.23      0.13     2.98     3.49 1.00     4255     3432
## day                  -0.07      0.01    -0.08    -0.05 1.00     9410     2849
## slphrs               -0.08      0.02    -0.11    -0.05 1.00     4483     3301
## erq_supp_c            0.09      0.10    -0.11     0.30 1.00     4046     2911
## slphrs:erq_supp_c    -0.01      0.01    -0.03     0.02 1.00     3761     3197
##
## 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     2622     2505
##
## 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).
```
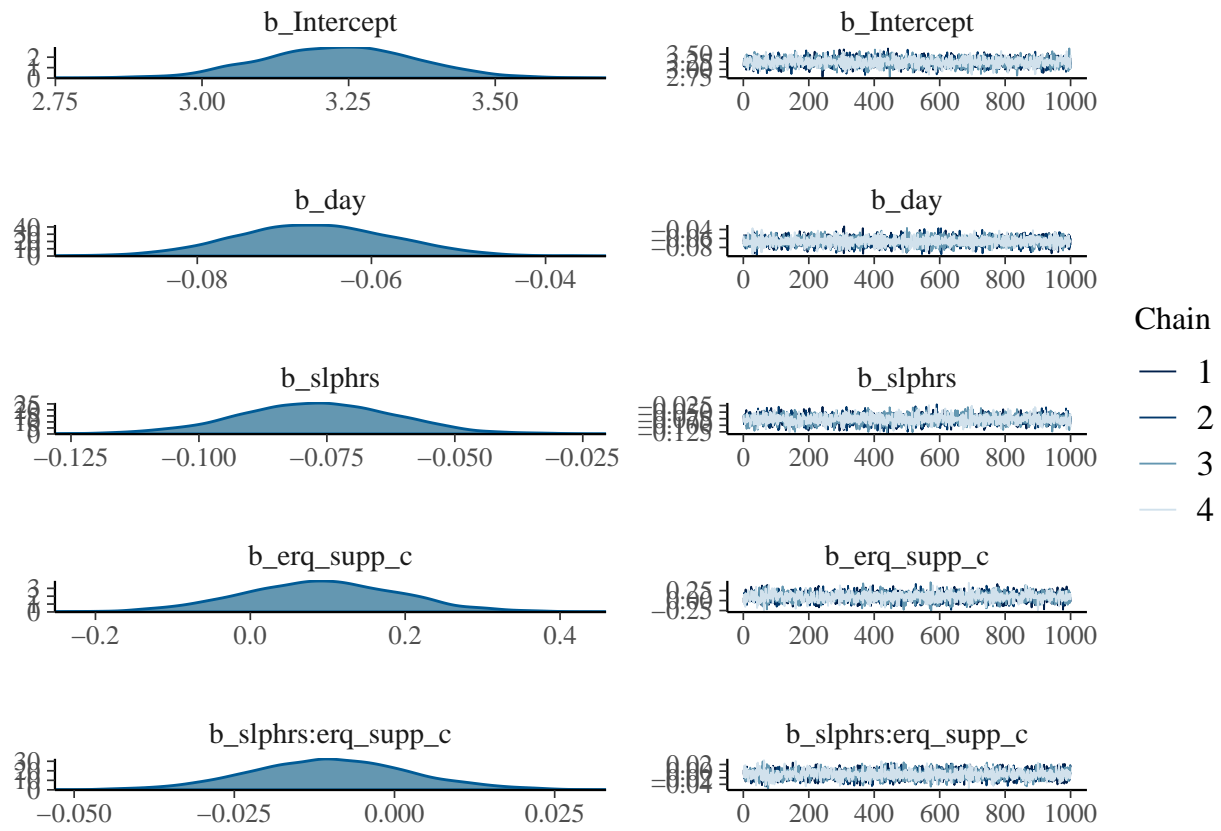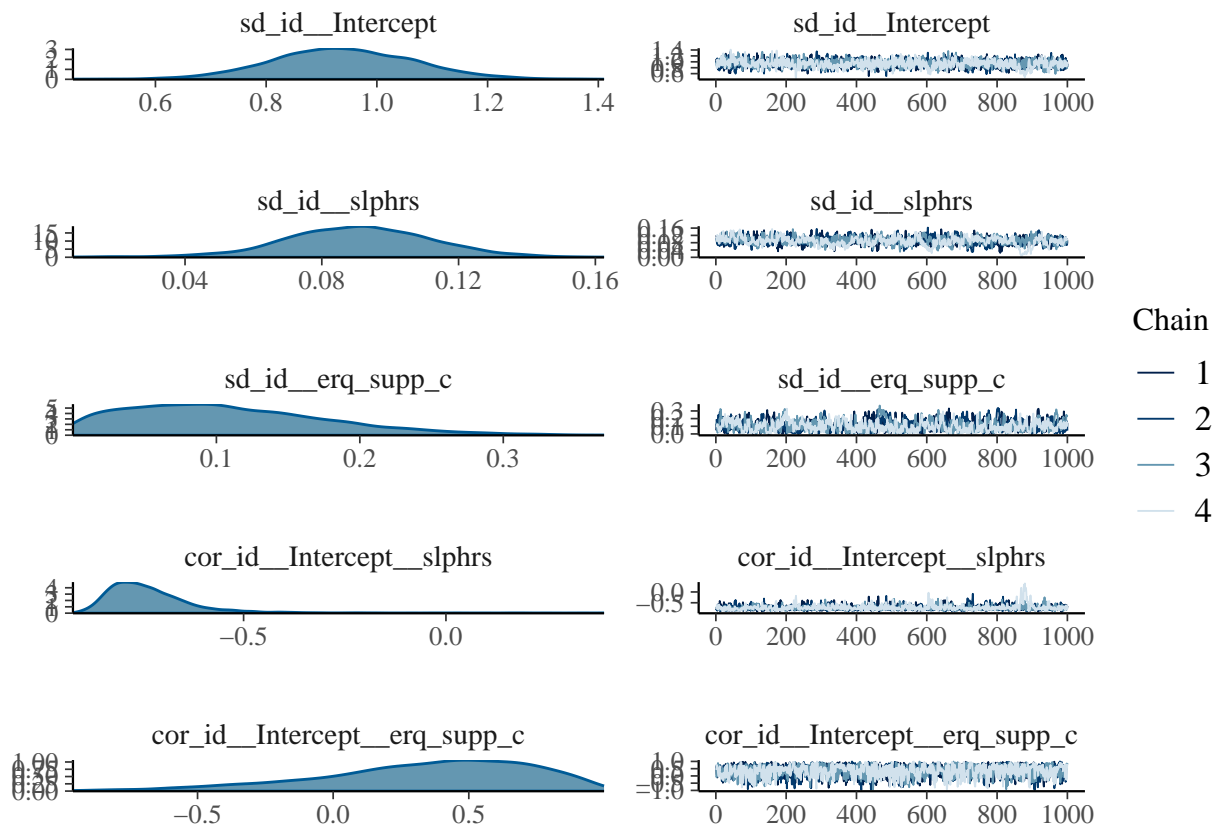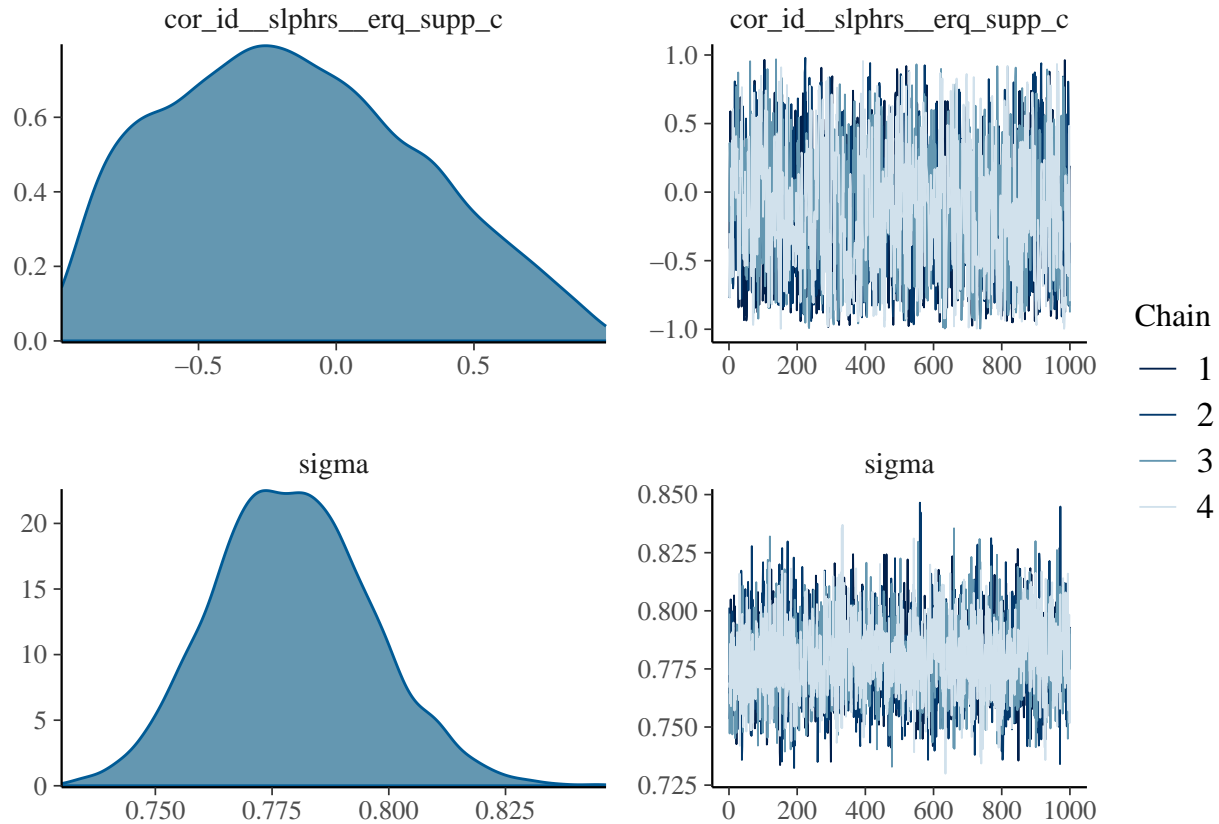
```r
bayes_R2(bpe.supp)
```

```
##    Estimate  Est.Error     Q2.5      Q97.5
## R2 0.443929 0.01789645 0.407381 0.4774714
```

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

TODO: interpret model results (including Bayesian significance testing) TODO: make a visualization of this model that shows the prototypical individual and all individuals

TODO: compare the two models (reap vs. supp - which one seems more predictive of better control of negaff?)

TODO: write up a final analysis/conclusion for this tutorial

Conclusion

[TODO]

References

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

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.