

Complete this during the 2nd slide show!

Hierarchical linear regression in brms

We will again use the PDM data set, but with some different variables. You can load it using the following code in R:

```
library(curl)

# See https://github.com/mdnunez/encodingN200 for more information about the data
pdm_dat <- curl("https://tinyurl.com/PDMdataESCP2022")
pdm <- read.csv(pdm_dat)

colnames(pdm) <- c('N200_latencies', 'N200_amplitudes',
  'RT', 'accuracy', 'condition', 'EEG_session',
  'experiment', 'session', 'subject')

pdm <- pdm[pdm$experiment == 1, ]

pdm$N200_latencies <- pdm$N200_latencies/1000

pdm$RT <- pdm$RT/1000

head(pdm)
```

1. (10min) What is the brms code to estimate a linear regression with *RT* as the dependent variable, *N200\_latencies* and *N200\_amplitudes* as the independent variables, and *an interaction term*?

The `summary()` should provide something like the output on the next page.

```

## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: RT ~ N200_latencies * N200_amplitudes
## Data: pdm (Number of observations: 5532)
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
## total post-warmup draws = 4000
##
## Population-Level Effects:
##
```

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat
Intercept	0.62	0.03	0.56	0.69	1.00
N200_latencies	0.83	0.15	0.53	1.14	1.00
N200_amplitudes	-0.01	0.02	-0.04	0.02	1.00
N200_latencies:N200_amplitudes	0.01	0.08	-0.15	0.17	1.00

```

## Bulk_ESS Tail_ESS
## Intercept 1324 1532
## N200_latencies 1374 1564
## N200_amplitudes 1186 1587
## N200_latencies:N200_amplitudes 1203 1603
##
## Family Specific Parameters:
## Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma 0.24 0.00 0.23 0.24 1.00 2569 2104
##
## 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).

```

## 2. (5min) What effects are significant in this model?

```
bayes_anova <- brm(RT ~ factor(condition)*factor(accuracy), data=pdm)
summary(bayes_anova)
```

```
summary(bayes_anova)
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: RT ~ factor(condition) * factor(accuracy)
## Data: pdm (Number of observations: 5532)
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
## total post-warmup draws = 4000
##
## Population-Level Effects:
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat
## Intercept	0.84	0.01	0.83	0.86	1.00
## factorcondition1	-0.05	0.01	-0.07	-0.02	1.00
## factorcondition2	-0.02	0.01	-0.04	0.01	1.00
## factoraccuracy1	-0.03	0.01	-0.06	-0.01	1.00
## factorcondition1:factoraccuracy1	-0.03	0.02	-0.06	0.00	1.01
## factorcondition2:factoraccuracy1	-0.06	0.02	-0.09	-0.02	1.00

```
## Bulk_ESS Tail_ESS
## Intercept 1831 2621
## factorcondition1 1804 2485
## factorcondition2 2082 2580
## factoraccuracy1 1500 2275
## factorcondition1:factoraccuracy1 1431 2305
## factorcondition2:factoraccuracy1 1801 2071
##
## Family Specific Parameters:
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma 0.24 0.00 0.23 0.24 1.00 3539 2794
##
## 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).
```

3. (10min) What is the brms code to estimate a linear regression with *RT* as the dependent variable, *N200\_latencies* and *N200\_amplitudes* as the independent variables, an *interaction term*, and a random intercept for each *subject*?

The summary() should output something like this:

```
## Warning: There were 1 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: RT ~ (1 | subject) + N200_latencies * N200_amplitudes
## Data: pdm (Number of observations: 5532)
## Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
## total post-warmup draws = 4000
##
## Group-Level Effects:
## ~subject (Number of levels: 12)
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)    0.13     0.03    0.09    0.21 1.01      774    1318
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat
## Intercept           0.61     0.05    0.52    0.71 1.01
## N200_latencies       0.86     0.14    0.59    1.13 1.00
## N200_amplitudes      0.00     0.02   -0.03    0.03 1.00
## N200_latencies:N200_amplitudes -0.02    0.07   -0.16    0.12 1.00
##           Bulk_ESS Tail_ESS
## Intercept           550    1359
## N200_latencies      1544    1699
## N200_amplitudes     1423    1693
## N200_latencies:N200_amplitudes 1396    1548
##
## Family Specific Parameters:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma       0.21     0.00    0.21    0.22 1.00     2406    2304
##
## 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).
```

