Momentary Subjective Happiness Analysis

Groundhog Day Project

Corrado Caudek

2024-08-10

Prelims

```
# Load necessary library
suppressPackageStartupMessages({
    library("here")
    library("tidyverse")
    library("mice")
    library("lme4")
    library("brms")
    library("bayesplot")
    library("effectsize")
    library("scales")
    library("sjstats")
    library("sjPlot")
    library("sjmisc")
})
```

Source functions

```
source(
    here::here("workflows", "scripts", "funs", "funs_instant_mood.R")
)
```

Import Data

```
# This script requires the dataframe d from prl_mood.R.
d1 <- readRDS(
    here::here(
        "data", "prep", "groundhog_all_clean.RDS"
    )
)

d1$user_id <- as.numeric(d1$user_id)

# Remove last ema sessions because compliance is too low
d <- d1 |>
    dplyr::filter(!is.na(user_id) & ema_number < 13)

length(unique(d$user_id))
```

Data Wrangling

```
set.seed(123)

# Standardize momentary mood by user_id.
dz <- d |>
    group_by(user_id) |>
    mutate(
      zim = as.vector(scale(instant_mood, center = TRUE, scale = TRUE))
    ) |>
    ungroup()

# Remove participants because of convergence problems.
bad_id_indices <- c(93, 99, 127, 146, 195, 216)
user_id_codes <- unique(dz$user_id)

bad_codes <- c(
    user_id_codes[bad_id_indices],
    3338029881, 3665345709, 3248648540
)

good_codes <- setdiff(user_id_codes, bad_codes)
dz_clean <- dz[dz$user_id %in% good_codes, ]
```

```r
# Get list of unique user_ids
user_id_codes <- unique(dz_clean$user_id)
length(user_id_codes)
```

```
[1] 215
```

Model of Momentary Subjective Well-Being

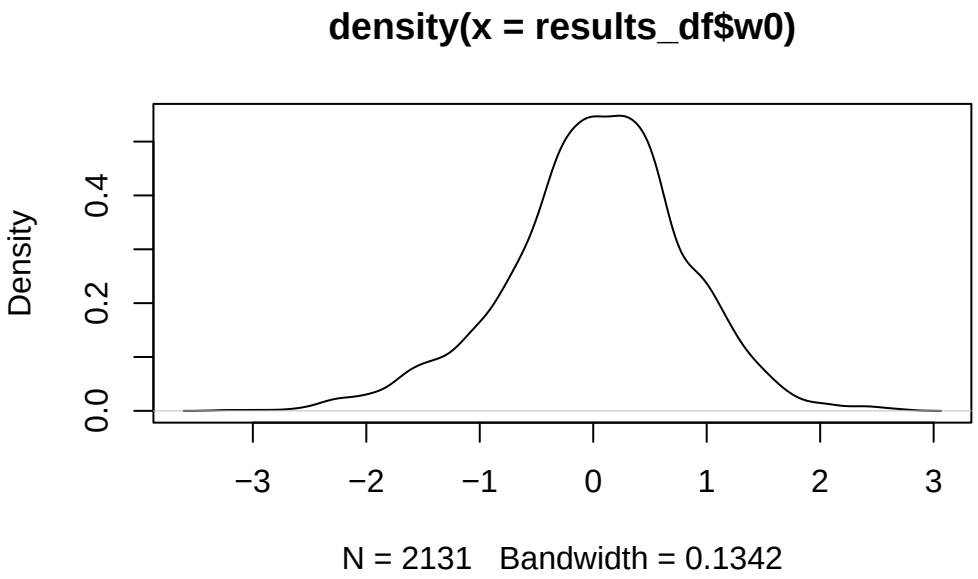
Estimation of the parameters of the Model of subjective well-being (ispired by **https://doi.org/10.1073/pnas.**

```r
# Apply the function process_user() to each user_id.
par_list <- NULL
results_list <- lapply(user_id_codes, process_user)
```

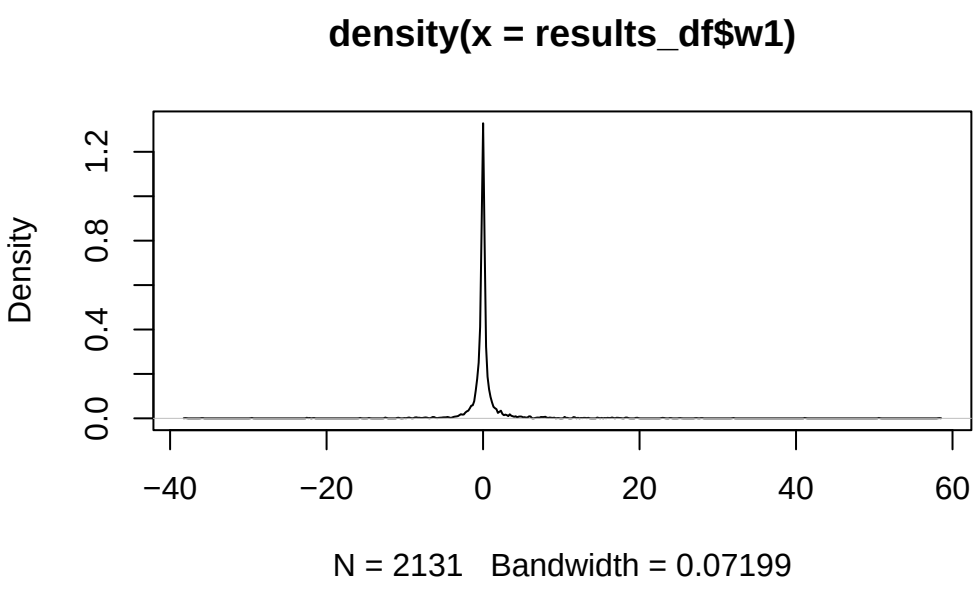
```r
# Bind all data frames together into a single data frame
all_results_df <- bind_rows(results_list)

# Add mood_pre, mood_post, control.
bysubj_mood_df <- dz_clean |>
    group_by(user_id, ema_number) |>
    summarize(
        mood_pre = mean(mood_pre),
        mood_post = mean(mood_post),
        control = mean(control)
    ) |>
    ungroup()

# Final dataframe.
results_df <- left_join(
    all_results_df, bysubj_mood_df,
    by = c("user_id", "ema_number")
)

names(results_df)
```

Examine and Clean Parameters

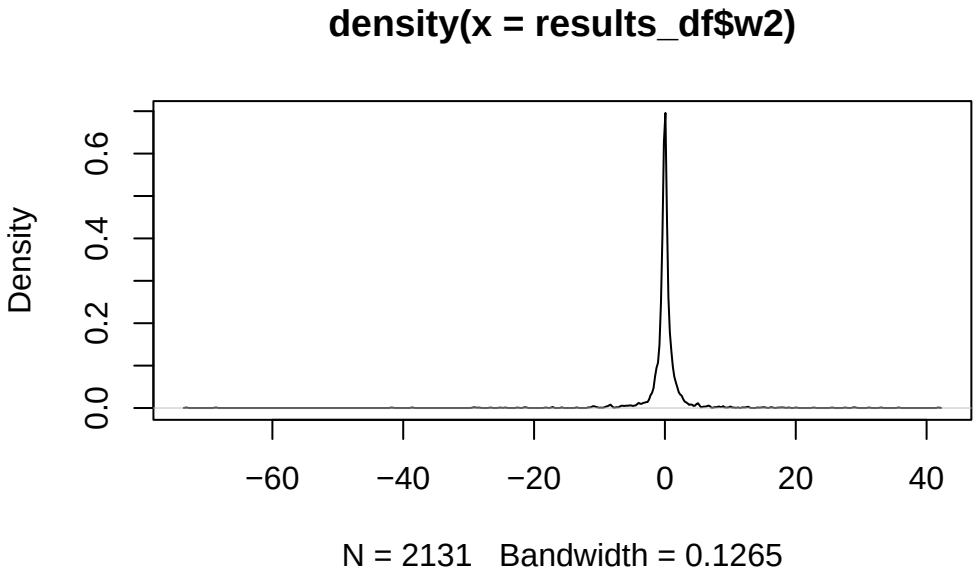
```
plot(density(results_df$w0))
```



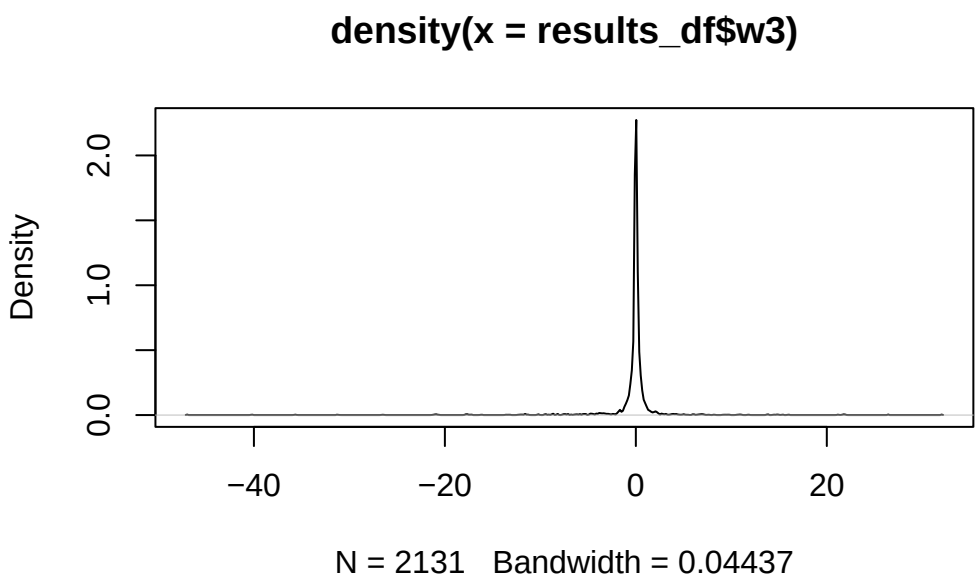

```
plot(density(results_df$w1))
```



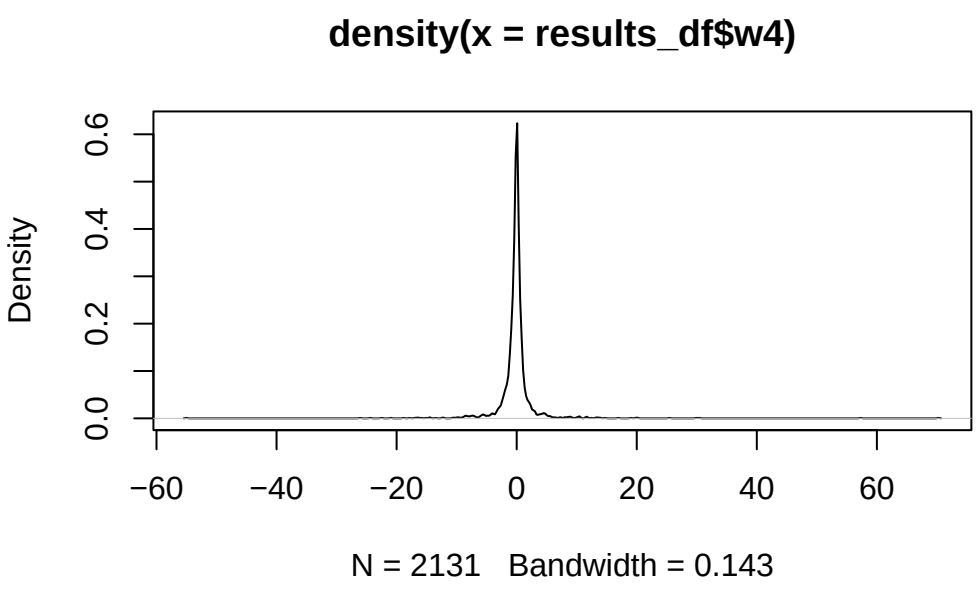

```
plot(density(results_df$w2))
```



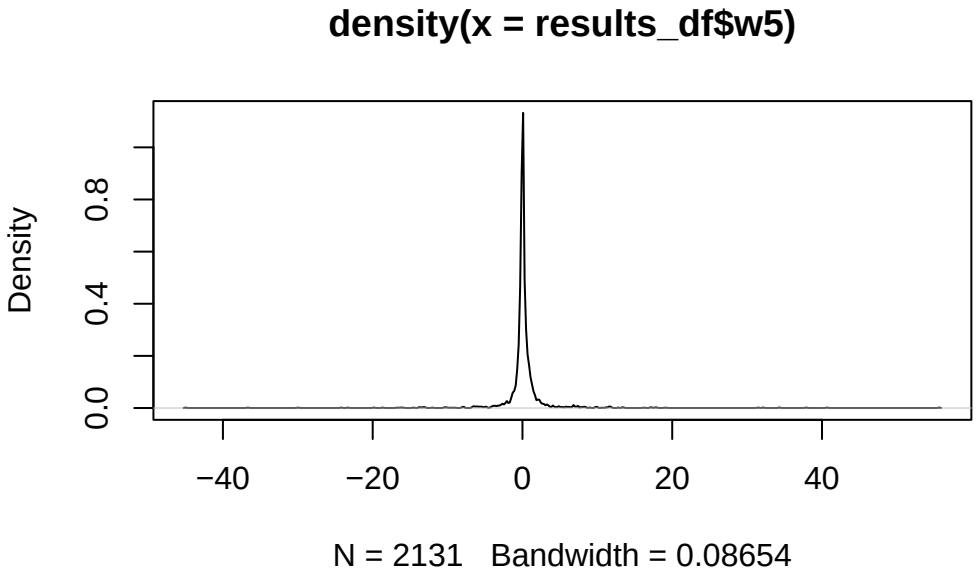
```
plot(density(results_df$w3))
```



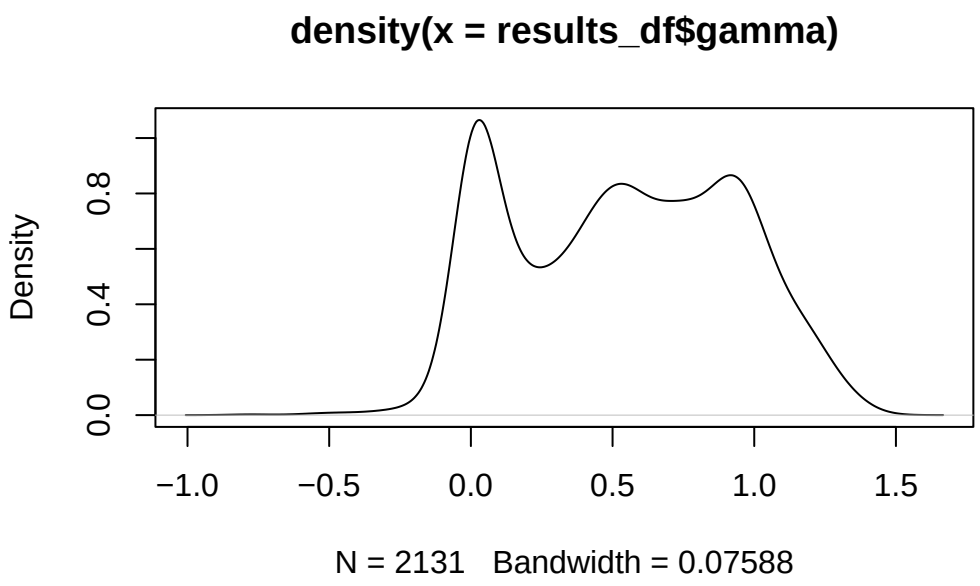

```
plot(density(results_df$w4))
```



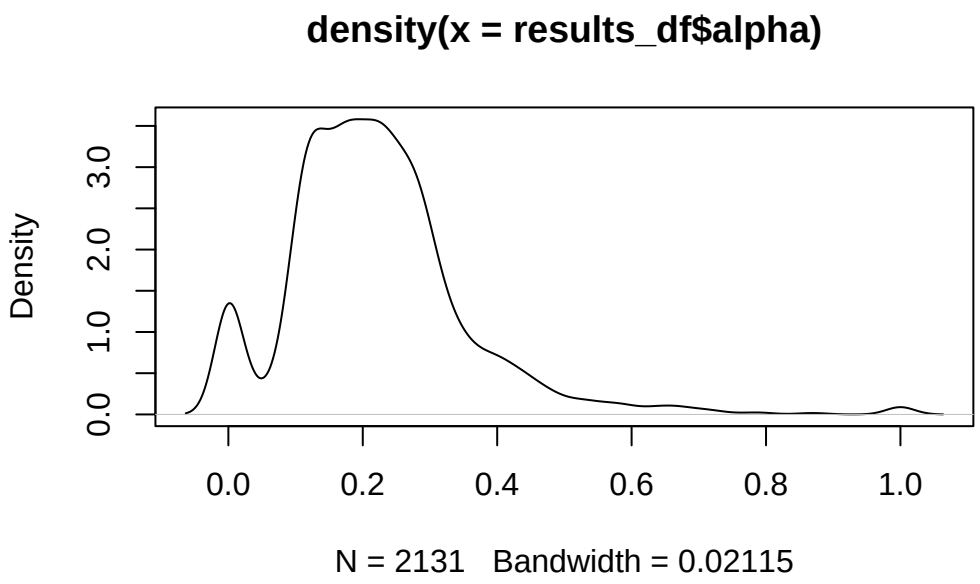

```
plot(density(results_df$w5))
```



```r
plot(density(results_df$gamma))
```




```r
plot(density(results_df$alpha))
```




```r
remove_outliers <- function(my_vector) {
    # Calculate mean and standard deviation
    mean_val <- mean(my_vector, na.rm = TRUE)
    std_dev <- sd(my_vector, na.rm = TRUE)
    # Calculate Z-scores
    z_scores <- (my_vector - mean_val) / std_dev
    # Replace outliers with NA
    my_vector[abs(z_scores) > 3.5] <- NA
    my_vector
}

results_df$cw1 <- remove_outliers(results_df$w1)
results_df$cw2 <- remove_outliers(results_df$w2)
results_df$cw3 <- remove_outliers(results_df$w3)
results_df$cw4 <- remove_outliers(results_df$w4)
results_df$cw5 <- remove_outliers(results_df$w5)
results_df$cwg <- remove_outliers(results_df$gamma)

results_df$cmood_pre <- remove_outliers(results_df$mood_pre)
results_df$cmood_post <- remove_outliers(results_df$mood_post)
```

```r
# Remove and impute the values alpha = 0 ad alpha = 1
results_df$calpha <- ifelse(
  results_df$alpha < 0.00001 | results_df$alpha > 0.99999, NA, results_df$alpha
  )

# Remove redundant columns
results_w_df <- results_df |>
    dplyr::select(
      !c(w0, w1, w2, w3, w4, w5, gamma, mood_pre, mood_post, alpha)
    )
```

Imputation of missing data.

```r
imputed_cart <- complete(mice(results_w_df, method = "cart"))
```

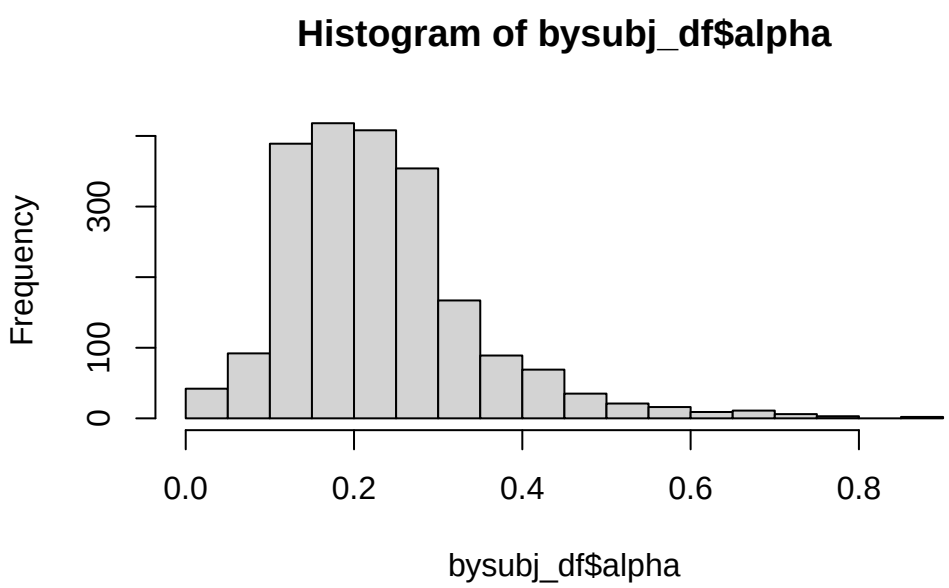
```
 iter imp variable
  1   1  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  1   2  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  1   3  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  1   4  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  1   5  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  2   1  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  2   2  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  2   3  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  2   4  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  2   5  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  3   1  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  3   2  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  3   3  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  3   4  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  3   5  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  4   1  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  4   2  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  4   3  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  4   4  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  4   5  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  5   1  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  5   2  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  5   3  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  5   4  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
  5   5  cw1  cw2  cw3  cw4  cw5  cmood_post  calpha
```

Data wrangling

```
# Difference post - pre on the raw mood data.
bysubj_df <- imputed_cart |>
    mutate(
        # Approximately centered:
        ema_number_c = ema_number - 4,
        # The mood_pre and mood_post outliers have been replaced
        # with NAs and then imputed.
        mood_dif = cmood_post - cmood_pre
    )

bysubj_df_renamed <- dplyr::rename(
  bysubj_df,
  mood_pre = cmood_pre,
  mood_post = cmood_post,
  alpha = calpha
)

bysubj_df$environment <- ifelse(
  bysubj_df$is_reversal == 1, "Volatile", "Stable"
  )

bysubj_df$alpha <- bysubj_df$calpha
bysubj_df$calpha <- NULL
```

Effect of Environment on Alpha

```
bysubj_df$alpha |> hist()
```




```
m <- brm(
    alpha ~ environment +
      (environment | user_id / ema_number),
    family = asym_laplace(),
    algorithm = "meanfield",
    data = bysubj_df,
    refresh = 0
)
```

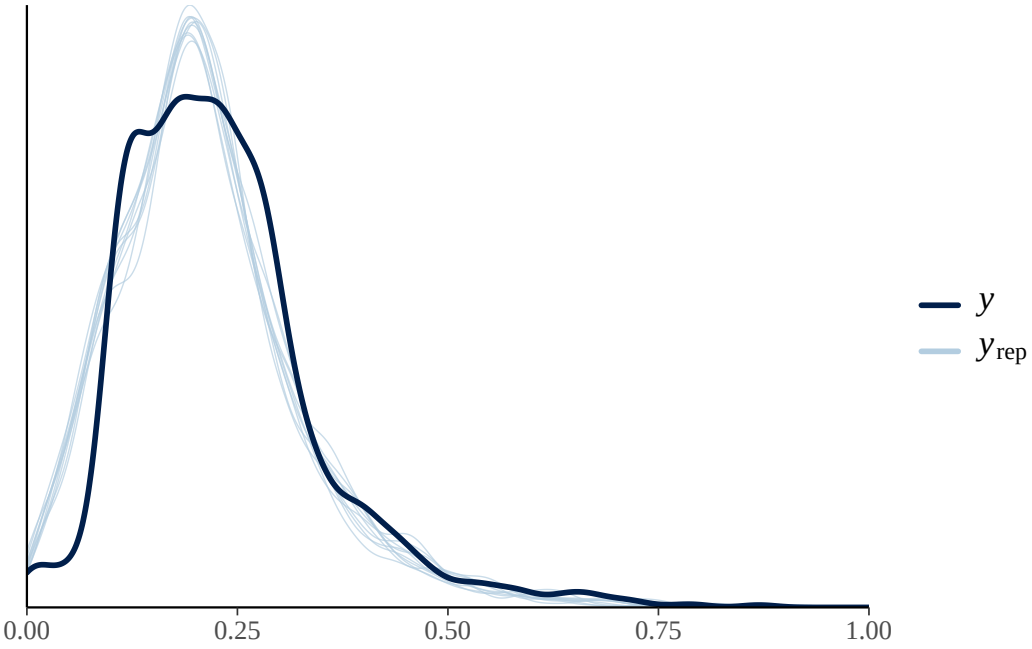
```
Running /Library/Frameworks/R.framework/Resources/bin/R CMD SHLIB foo.c
using C compiler: ‘Apple clang version 15.0.0 (clang-1500.0.40.1)’
using SDK: ‘MacOSX14.0.sdk’
clang -arch arm64 -I"/Library/Frameworks/R.framework/Resources/include" -DNDEBUG   -I"/Libra
In file included from <built-in>:1:
In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/St
In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/Rc
In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/Rc
/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/RcppEigen/include/Eigen/
namespace Eigen {
^
/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/RcppEigen/include/Eigen/
namespace Eigen {
```

```
                ^
                ;
In file included from <built-in>:1:
In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/St
In file included from /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/Rc
/Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/library/RcppEigen/include/Eigen/
#include <complex>
         ^~~~~~~~~
3 errors generated.
make: *** [foo.o] Error 1
```

```
pp_check(m) + xlim(0, 1)
```

```
Using 10 posterior draws for ppc type 'dens_overlay' by default.
```

```
Warning: Removed 552 rows containing non-finite values (`stat_density()`).
```




```
summary(m)
```

```
 Family: asym_laplace
  Links: mu = identity; sigma = identity; quantile = identity
Formula: alpha ~ environment + (environment | user_id/ema_number)
   Data: bysubj_df (Number of observations: 2131)
  Draws: 1 chains, each with iter = 1000; warmup = 0; thin = 1;
         total post-warmup draws = 1000

Group-Level Effects:
~user_id (Number of levels: 215)
                                  Estimate Est.Error l-95% CI u-95% CI Rhat
sd(Intercept)                         0.00      0.00     0.00     0.00 1.00
sd(environmentVolatile)               0.01      0.00     0.01     0.02 1.00
cor(Intercept,environmentVolatile)   -0.45      0.13    -0.68    -0.17 1.00
                                  Bulk_ESS Tail_ESS
sd(Intercept)                          882      787
sd(environmentVolatile)                937      941
cor(Intercept,environmentVolatile)     921     1026

~user_id:ema_number (Number of levels: 2131)
                                  Estimate Est.Error l-95% CI u-95% CI Rhat
sd(Intercept)                         0.00      0.00     0.00     0.01 1.00
sd(environmentVolatile)               0.00      0.00     0.00     0.01 1.00
cor(Intercept,environmentVolatile)   -0.05      0.50    -0.89     0.85 1.00
                                  Bulk_ESS Tail_ESS
sd(Intercept)                          906      952
sd(environmentVolatile)               1011      903
cor(Intercept,environmentVolatile)     977      981

Population-Level Effects:
                   Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept              0.09      0.01     0.08     0.10 1.00      975      914
environmentVolatile    0.10      0.01     0.09     0.11 1.00      940      981

Family Specific Parameters:
         Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma        0.03      0.00     0.03     0.04 1.00      933      990
quantile     0.36      0.01     0.34     0.37 1.00      770      856

Draws were sampled using variational(meanfield).
```

```
bayes_R2(m)
```

```
    Estimate  Est.Error      Q2.5     Q97.5
R2 0.1549301 0.01496497 0.1261506 0.1843055
```

```
conditional_effects(m, "environment")
```