final project codes

Minci Zhang

2022-12-07

```
# import data
dat<-read_sav (here ("data_files", "y1y2.sav"))
dat <- dat%>%
mutate(ethnic= as.factor(ifelse(ethnic1final ==1, 0, 1))) #recode ethnicity: H=0, A=1
```

```
# wide to long and person mean center
dat_long <- dat %>%
    pivot_longer(
        c(like1,like2, pop1, pop2, Gpa1,Gpa2),
        names_to = c(".value", "time"),
        names_pattern = "(like|pop|Gpa)([1-2])",
        names_transform = list(time = as.integer)
    )

dat_long <- dat_long %>%
    group_by(id) %>%
    mutate(Gpa_pm = mean(Gpa),
           Gpa_pmc = Gpa - Gpa_pm) %>%
    ungroup()

dat_long <- dat_long %>%
    mutate(time = time - 1)%>%
    select(id, ethnic, time,like, pop, Gpa, Gpa_pm, Gpa_pmc)
```

```
# descriptive stats full sample
 datasummary( (like1 + like2 + pop1 + pop2 + Gpa1 + Gpa2) ~
                (Mean + SD ),
            data = dat)
```

```
## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output
## %in% : 'length(x) = 2 > 1' in coercion to 'logical(1)'
```

	Mean	SD
like1	0.04	1.00

like2	0.06	0.97
pop1	−0.04	0.94
pop2	0.04	1.04
Gpa1	3.04	0.85
Gpa2	3.17	0.81

```
# descriptive stats by ethnicity
 datasummary( (like1 + like2 + pop1 + pop2 + Gpa1 + Gpa2)  *
                (Mean + SD ) ~
                factor(ethnic, labels = c("Hispanic", "Asian")),
            data = dat)
```

		Hispanic	Asian
like1	Mean	0.03	0.05
	SD	0.98	1.00
like2	Mean	−0.10	0.14
	SD	0.91	0.99
pop1	Mean	0.03	−0.08
	SD	0.93	0.95
pop2	Mean	−0.10	0.11
	SD	0.85	1.13
Gpa1	Mean	2.37	3.39
	SD	0.81	0.62
Gpa2	Mean	2.54	3.51
	SD	0.77	0.60

```
# correlation

 dat %>%
 select(like1, like2, pop1, pop2, Gpa2, Gpa2) %>%
  datasummary_correlation(method="pearson")
```

	like1	like2	pop1	pop2	Gpa2
like1	1	.	.	.	.
like2	.41	1	.	.	.
pop1	.51	.30	1	.	.
pop2	.31	.42	.59	1	.
Gpa2	.11	.20	−.08	.03	1

```
# correlation by ethnic groups
hisp<- dat %>% filter(ethnic== 0)
asian <- dat %>% filter(ethnic == 1)


hisp %>%
  select(like1, like2, pop1, pop2, Gpa2, Gpa2) %>%
  datasummary_correlation(method="pearson")
```

	like1	like2	pop1	pop2	Gpa2
like1	1	.	.	.	.
like2	.43	1	.	.	.
pop1	.55	.28	1	.	.
pop2	.50	.51	.59	1	.
Gpa2	.19	.19	−.14	−.02	1

```
asian %>%
  select(like1, like2, pop1, pop2, Gpa2, Gpa2) %>%
  datasummary_correlation(method="pearson")
```

	like1	like2	pop1	pop2	Gpa2
like1	1	.	.	.	.
like2	.40	1	.	.	.
pop1	.49	.32	1	.	.
pop2	.24	.38	.60	1	.
Gpa2	.08	.15	.00	−.04	1

social accaptance analyses

##equations:

main effects

$$\begin{aligned}\text{soc_acc}_{ti} = {} & \gamma_{00} + \gamma_{01}\text{ethnicity}_{ti} + \gamma_{02}\text{GPA_pm}_{ti} + \gamma_{10}\text{time}_{ti} + \gamma_{20}\text{GPA_pmc}_{ti} + \\ & u_{1i}\text{time}_{ti} + u_{2i}\text{GPA_pmc}_{ti} + \\ & u_{0i} + e_{ti}\end{aligned}$$

2-way interactions

$$\begin{aligned}\text{soc_acc}_{ti} = {} & \gamma_{00} + \gamma_{01}\text{ethnicity}_{ti} + \gamma_{02}\text{GPA_pm}_{ti} + \gamma_{10}\text{time}_{ti} + \gamma_{20}\text{GPA_pmc}_{ti} + \\ & \gamma_{03}\text{ethnicity}_{ti} \times \text{GPA_pm}_{ti} + \gamma_{11}\text{ethnicity}_{ti} \times \text{time}_{ti} + \\ & \gamma_{12}\text{GPA_pm}_{ti} \times \text{time}_{ti} + \gamma_{21}\text{ethnicity}_{ti} \times \text{GPA_pmc}_{ti} + \\ & \gamma_{30}\text{time} \times \text{GPA_pmc}_{ti} + \\ & u_{1i}\text{time}_{ti} + u_{0i} + e_{ti}\end{aligned}$$

3-way interactions

$$

$$\begin{aligned}\text{soc_acc}_{ti} = {} & \gamma_{00} + \gamma_{01}\text{ethnicity}_{ti} + \gamma_{02}\text{GPA_pm}_{ti} + \gamma_{10}\text{time}_{ti} + \gamma_{20}\text{GPA_pmc}_{ti} + \\ & \gamma_{03}\text{ethnicity}_{ti} \times \text{GPA_pm}_{ti} + \gamma_{11}\text{ethnicity}_{ti} \times \text{time}_{ti} + \\ & \gamma_{12}\text{GPA_pm}_{ti} \times \text{time}_{ti} + \gamma_{21}\text{ethnicity}_{ti} \times \text{GPA_pmc}_{ti} + \\ & \gamma_{30}\text{time} \times \text{GPA_pmc}_{ti} + \\ & \gamma_{31}\text{ethnicity}_{ti} \times \text{time} \times \text{GPA_pmc}_{ti} + \gamma_{13}\text{ethnicity}_{ti} \times \text{GPA_pm}_{ti} \times \text{time}_{ti} + \\ & u_{1i}\text{time}_{ti} + u_{0i} + e_{ti}\end{aligned}$$

$$

ICC

```
# fit a random intercept model
m0<- lmer(like ~  1 + (1|id), data=dat_long)
summary(m0)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: like ~ 1 + (1 | id)
##    Data: dat_long
##
## REML criterion at convergence: 1789.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.9819 -0.6111 -0.1127  0.4905  2.5298
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept) 0.3893   0.6239
##  Residual             0.5722   0.7564
## Number of obs: 659, groups:  id, 335
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)   0.05304    0.04512 335.27000   1.175    0.241
```

```
ICC<- 0.3893/ (0.3893+ 0.5722 )
ICC
```

```
## [1] 0.4048882
```

model1: main effects

```
#test random slopes
# a model with only random slope of time
m1 <- brm(bf(like ~ Gpa_pm+ Gpa_pmc + time + ethnic + (time|id),sigma ~ 0 ), data=dat
_long,
             seed = 1150,
             file = "final1",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc (in addition to time)
m2 <- brm(bf(like ~ Gpa_pm+ Gpa_pmc + time + ethnic + (Gpa_pmc + time|id),sigma ~ 0
), data=dat_long,
             seed = 1151,
             file = "final2",
             chains = 2L, iter = 1000L)

loo(m1, m2)
```

```
## Output of model 'm1':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -900.8 13.3
## p_loo        58.7  2.9
## looic      1801.6 26.7
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     651   98.8%   211
##  (0.5, 0.7]   (ok)         8    1.2%   368
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm2':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -900.7 13.2
## p_loo        60.8  3.0
## looic      1801.4 26.4
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     640   97.1%   254
##  (0.5, 0.7]   (ok)        18    2.7%   283
##    (0.7, 1]   (bad)        1    0.2%   699
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##    elpd_diff se_diff
## m2  0.0       0.0
## m1 -0.1       0.5
```

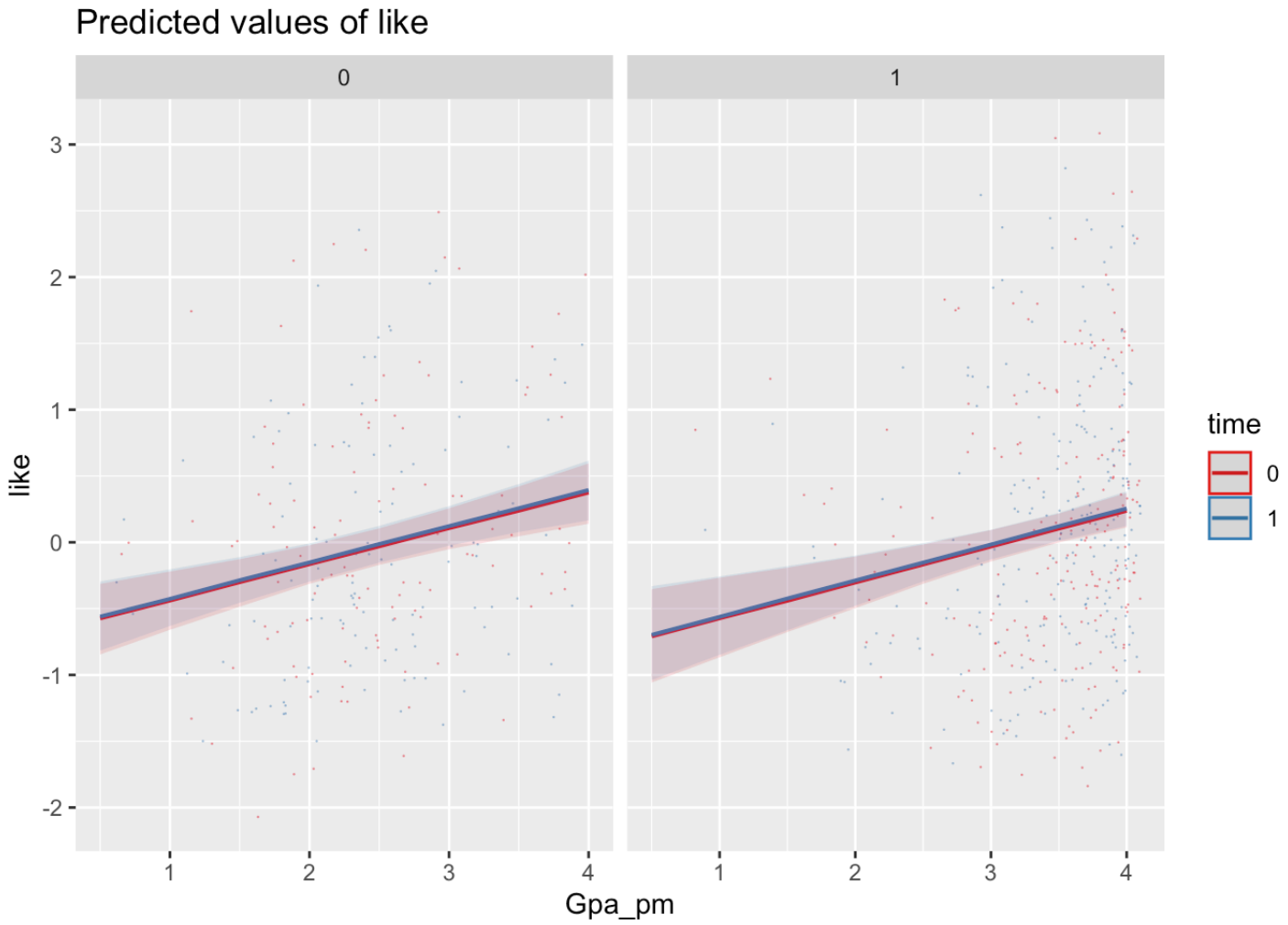
```
# m2 has smaller looic value.
#  include random slope of time and GPA_pmc (opt to m2)


#model2 results
summary(m2)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = log
## Formula: like ~ Gpa_pm + Gpa_pmc + time + ethnic + (Gpa_pmc + time | id)
##          sigma ~ 0
##    Data: dat_long (Number of observations: 659)
##   Draws: 2 chains, each with iter = 1000; warmup = 500; thin = 1;
##          total post-warmup draws = 1000
##
## Group-Level Effects:
## ~id (Number of levels: 335)
##                       Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)             0.38      0.08     0.22     0.52 1.01      391
## sd(Gpa_pmc)               0.16      0.12     0.01     0.46 1.00      859
## sd(time)                  0.10      0.07     0.00     0.26 1.00      353
## cor(Intercept,Gpa_pmc)    0.02      0.52    -0.87     0.89 1.00     1538
## cor(Intercept,time)      -0.11      0.48    -0.90     0.81 1.00      929
## cor(Gpa_pmc,time)        -0.02      0.51    -0.89     0.88 1.00      596
##                       Tail_ESS
## sd(Intercept)              508
## sd(Gpa_pmc)                498
## sd(time)                   340
## cor(Intercept,Gpa_pmc)     556
## cor(Intercept,time)        770
## cor(Gpa_pmc,time)          632
##
## Population-Level Effects:
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept    -0.71      0.20    -1.09    -0.32 1.00     1052      805
## Gpa_pm        0.27      0.07     0.13     0.41 1.00      924      685
## Gpa_pmc      -0.02      0.15    -0.32     0.25 1.00     1939      868
## time          0.02      0.08    -0.13     0.17 1.00     2826      782
## ethnic1      -0.14      0.11    -0.36     0.10 1.00      935      730
##
## 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).
```

```
tab_model(m2)
```

	like	
Predictors	*Estimates*	*CI (95%)*
Intercept	-0.71	-1.09 – -0.32
Gpa_pm	0.27	0.13 – 0.41
Gpa_pmc	-0.02	-0.32 – 0.25
time	0.01	-0.13 – 0.17
ethnic: ethnic1	-0.14	-0.36 – 0.10
Random Effects		
σ^2	1.00	
$\tau_{00\ id}$	0.15	
$\tau_{11\ id.Gpa_pmc}$	0.04	
$\tau_{11\ id.time}$	0.01	
ρ_{01}		
ρ_{01}		
ICC	0.13	
N_{id}	335	
Observations	659	
Marginal R^2 / Conditional R^2	0.039 / 0.193	

```
plot_model(m2,
    type = "pred", terms = c("Gpa_pm", "time", "ethnic"),
    show.data = TRUE, jitter = 0.1,
    dot.alpha = 0.5, dot.size = 0.1,
)
```



model2: two way interactions:

```
#test random slopes
# a model with only random slope of time
m3 <- brm(bf(like ~ (Gpa_pm+ Gpa_pmc + time ) * ethnic + Gpa_pmc * time + Gpa_pm * ti
me + (time|id),sigma ~ 0 ), data=dat_long,
             seed = 1150,
             file = "final3",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc (in addition to time)
m4 <- brm(bf(like ~ (Gpa_pm+ Gpa_pmc + time ) * ethnic + Gpa_pmc * time + Gpa_pm * ti
me + (Gpa_pmc + time|id),sigma ~ 0 ), data=dat_long,
             seed = 1151,
             file = "final4",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc * time
m5 <- brm(bf(like ~ (Gpa_pm+ Gpa_pmc + time ) * ethnic + Gpa_pmc * time + Gpa_pm * ti
me + (Gpa_pmc * time|id),sigma ~ 0 ), data=dat_long,
             seed = 1151,
             file = "final5",
             chains = 2L, iter = 1000L)



loo(m3, m4, m5)
```

```
## Output of model 'm3':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -902.8 13.3
## p_loo        62.6  3.1
## looic      1805.6 26.6
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     646   98.0%   266
##  (0.5, 0.7]   (ok)        13    2.0%   710
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
```

```
##
## Output of model 'm4':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -903.1 13.3
## p_loo        63.0  3.1
## looic      1806.2 26.5
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     651   98.8%   230
##  (0.5, 0.7]   (ok)         8    1.2%   306
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm5':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -905.2 13.3
## p_loo        65.1  3.1
## looic      1810.3 26.6
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     640   97.1%   201
##  (0.5, 0.7]   (ok)        19    2.9%   373
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##    elpd_diff se_diff
## m3  0.0       0.0
## m4 -0.3       0.5
```

```
## m5 -2.4       0.5
```

```
# m3 has smaller looic value.
# only include random slope of time (opt to m3)


#model results
summary(m3)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = log
## Formula: like ~ (Gpa_pm + Gpa_pmc + time) * ethnic + Gpa_pmc * time + Gpa_pm * tim
e + (time | id)
##          sigma ~ 0
##    Data: dat_long (Number of observations: 659)
##   Draws: 2 chains, each with iter = 1000; warmup = 500; thin = 1;
##          total post-warmup draws = 1000
##
## Group-Level Effects:
## ~id (Number of levels: 335)
##                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)           0.38      0.07     0.24     0.53 1.01      292      592
## sd(time)                0.09      0.07     0.00     0.25 1.00      434      305
## cor(Intercept,time)    -0.10      0.58    -0.96     0.92 1.00      686      574
##
## Population-Level Effects:
##                 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept          -0.53      0.32    -1.16     0.13 1.00      552      523
## Gpa_pm              0.24      0.12     0.01     0.48 1.00      532      708
## Gpa_pmc             0.27      0.28    -0.25     0.81 1.00      820      812
## time               -0.25      0.34    -0.92     0.44 1.00      669      727
## ethnic1            -0.25      0.44    -1.09     0.58 1.00      498      506
## Gpa_pm:ethnic1      0.01      0.14    -0.25     0.28 1.00      459      430
## Gpa_pmc:ethnic1    -0.12      0.30    -0.72     0.49 1.00     1202      845
## time:ethnic1        0.19      0.21    -0.21     0.61 1.00     1120      717
## Gpa_pmc:time       -0.43      0.34    -1.06     0.25 1.00      863      735
## Gpa_pm:time         0.05      0.13    -0.20     0.29 1.00      565      760
##
## 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).
```

model3: 3 way interaction

```
#test random slopes
# a model with only random slope of time
m6 <- brm(bf(like ~ (Gpa_pm+ Gpa_pmc) * time * ethnic + (time|id),sigma ~ 0 ), data=d
at_long,
             seed = 1150,
             file = "final6",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc (in addition to time)
m7 <- brm(bf(like ~ (Gpa_pm+ Gpa_pmc) * time * ethnic + (Gpa_pmc + time|id),sigma ~ 0
), data=dat_long,
             seed = 1151,
             file = "final7",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc * time
m8 <- brm(bf(like ~ (Gpa_pm+ Gpa_pmc) * time * ethnic + (Gpa_pmc * time|id),sigma ~ 0
), data=dat_long,
             seed = 1151,
             file = "final8",
             chains = 2L, iter = 1000L)



loo(m6, m7, m8)
```

```
## Output of model 'm6':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -903.3 13.2
## p_loo        63.3  3.0
## looic      1806.5 26.5
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     644   97.7%   280
##  (0.5, 0.7]   (ok)        15    2.3%   303
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
```

```
##
## Output of model 'm7':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -905.1 13.3
## p_loo        64.0  3.1
## looic      1810.2 26.7
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     644   97.7%   237
##  (0.5, 0.7]   (ok)        15    2.3%   277
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm8':
##
## Computed from 1000 by 659 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -904.3 13.2
## p_loo        64.7  3.1
## looic      1808.7 26.5
## ------
## Monte Carlo SE of elpd_loo is 0.3.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     629   95.4%   270
##  (0.5, 0.7]   (ok)        30    4.6%   293
##    (0.7, 1]   (bad)        0    0.0%   <NA>
##    (1, Inf)   (very bad)   0    0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##    elpd_diff se_diff
## m6  0.0       0.0
## m8 -1.1       0.5
```

```
## m7 -1.9       0.5
```

```
# m6 has smaller looic value.
#  include random slope of time (opt to m6)


#model results
summary(m6)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = log
## Formula: like ~ (Gpa_pm + Gpa_pmc) * time * ethnic + (time | id)
##          sigma ~ 0
##    Data: dat_long (Number of observations: 659)
##   Draws: 2 chains, each with iter = 1000; warmup = 500; thin = 1;
##          total post-warmup draws = 1000
##
## Group-Level Effects:
## ~id (Number of levels: 335)
##                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)           0.38      0.08     0.23     0.54 1.00      358      513
## sd(time)                0.09      0.07     0.00     0.25 1.01      462      727
## cor(Intercept,time)    -0.14      0.58    -0.96     0.93 1.01      916      722
##
## Population-Level Effects:
##                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## Intercept               -0.64      0.32    -1.23     0.03 1.00      397
## Gpa_pm                   0.28      0.13     0.02     0.52 1.00      378
## Gpa_pmc                  0.07      0.32    -0.57     0.70 1.00      560
## time                    -0.08      0.45    -0.95     0.78 1.00      337
## ethnic1                  0.02      0.52    -0.96     1.00 1.00      401
## Gpa_pm:time             -0.02      0.18    -0.36     0.31 1.00      313
## Gpa_pmc:time            -0.05      0.48    -0.97     0.89 1.01      537
## Gpa_pm:ethnic1          -0.08      0.17    -0.40     0.25 1.00      345
## Gpa_pmc:ethnic1          0.30      0.47    -0.57     1.20 1.00      555
## time:ethnic1            -0.25      0.71    -1.59     1.15 1.00      324
## Gpa_pm:time:ethnic1      0.15      0.23    -0.33     0.58 1.00      293
## Gpa_pmc:time:ethnic1    -0.78      0.69    -2.11     0.56 1.00      534
##                      Tail_ESS
## Intercept                 656
## Gpa_pm                    670
## Gpa_pmc                   743
## time                      626
## ethnic1                   603
## Gpa_pm:time               505
## Gpa_pmc:time              765
## Gpa_pm:ethnic1            608
## Gpa_pmc:ethnic1           770
## time:ethnic1              451
## Gpa_pm:time:ethnic1       527
## Gpa_pmc:time:ethnic1      850
##
## 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).
```

popularity analyses

equations

main effects

$$\begin{aligned}\text{pop}_{ti} = \gamma_{00} + \gamma_{01}\text{ethnicity}_{ti} + \gamma_{02}\text{GPA_pm}_{ti} + \gamma_{10}\text{time}_{ti} + \gamma_{20}\text{GPA_pmc}_{ti} + \\ u_{1i}\text{time}_{ti} + u_{0i} + e_{ti}\end{aligned}$$

2-way interactions

$$\begin{aligned}\text{pop}_{ti} = \gamma_{00} + \gamma_{01}\text{ethnicity}_{ti} + \gamma_{02}\text{GPA_pm}_{ti} + \gamma_{10}\text{time}_{ti} + \gamma_{20}\text{GPA_pmc}_{ti} + \\ \gamma_{03}\text{ethnicity}_{ti} \times \text{GPA_pm}_{ti} + \gamma_{11}\text{ethnicity}_{ti} \times \text{time}_{ti} + \\ \gamma_{12}\text{GPA_pm}_{ti} \times \text{time}_{ti} + \gamma_{21}\text{ethnicity}_{ti} \times \text{GPA_pmc}_{ti} + \\ \gamma_{30}\text{time} \times \text{GPA_pmc}_{ti} + \\ u_{3i}\text{time}_{ti} \times \text{GPA_pmc}_{ti} + u_{0i} + e_{ti}\end{aligned}$$

3-way interactions

$$

$$\begin{aligned}\text{pop}_{ti} = \gamma_{00} + \gamma_{01}\text{ethnicity}_{ti} + \gamma_{02}\text{GPA_pm}_{ti} + \gamma_{10}\text{time}_{ti} + \gamma_{20}\text{GPA_pmc}_{ti} + \\ \gamma_{03}\text{ethnicity}_{ti} \times \text{GPA_pm}_{ti} + \gamma_{11}\text{ethnicity}_{ti} \times \text{time}_{ti} + \\ \gamma_{12}\text{GPA_pm}_{ti} \times \text{time}_{ti} + \gamma_{21}\text{ethnicity}_{ti} \times \text{GPA_pmc}_{ti} + \\ \gamma_{30}\text{time} \times \text{GPA_pmc}_{ti} + \\ \gamma_{31}\text{ethnicity}_{ti} \times \text{time} \times \text{GPA_pmc}_{ti} + \gamma_{13}\text{ethnicity}_{ti} \times \text{GPA_pm}_{ti} \times \text{time}_{ti} + \\ u_{1i}\text{time}_{ti} + u_{2i}\text{GPA_pmc}_{ti} + u_{0i} + e_{ti}\end{aligned}$$

$$

ICC

```
# fit a random intercept model
m<- lmer(pop ~  1 + (1|id), data=dat_long)
summary(m)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: pop ~ 1 + (1 | id)
##    Data: dat_long
##
## REML criterion at convergence: 1739
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.2774 -0.4131 -0.1743  0.2805  3.7182
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept) 0.6151   0.7843
##  Residual             0.3947   0.6282
## Number of obs: 663, groups:  id, 335
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept) 9.259e-03  4.935e-02 3.240e+02   0.188    0.851
```

```
ICC<- 0.6151/ (0.6151+ 0.3947 )
ICC
```

```
## [1] 0.6091305
```

model1: main effects

```
#test random slopes
# a model with only random slope of time
m9 <- brm(bf(pop ~ Gpa_pm+ Gpa_pmc + time + ethnic + (time|id),sigma ~ 0 ), data=dat_
long,
             seed = 1150,
             file = "final9",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc (in addition to time)
m10 <- brm(bf(pop ~ Gpa_pm+ Gpa_pmc + time + ethnic + (Gpa_pmc + time|id),sigma ~ 0
), data=dat_long,
             seed = 1151,
             file = "final10",
             chains = 2L, iter = 1000L)

loo(m9, m10)
```

```
## Output of model 'm9':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -893.7 20.1
## p_loo        83.4  7.2
## looic      1787.4 40.2
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     610   92.0%   105
##  (0.5, 0.7]   (ok)        48    7.2%   86
##    (0.7, 1]   (bad)        5    0.8%   12
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm10':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -894.6 20.1
## p_loo        85.3  7.4
## looic      1789.2 40.2
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     635   95.8%   90
##  (0.5, 0.7]   (ok)        24    3.6%   70
##    (0.7, 1]   (bad)        4    0.6%   29
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##     elpd_diff se_diff
## m9   0.0       0.0
## m10 -0.9       0.9
```

```
# m9 has smaller looic value.
# only include random slope of time  (opt to m9)


#model2 results
summary(m9)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = log
## Formula: pop ~ Gpa_pm + Gpa_pmc + time + ethnic + (time | id)
##          sigma ~ 0
##    Data: dat_long (Number of observations: 663)
##   Draws: 2 chains, each with iter = 1000; warmup = 500; thin = 1;
##          total post-warmup draws = 1000
##
## Group-Level Effects:
## ~id (Number of levels: 335)
##                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)           0.51      0.07     0.38     0.64 1.01      416      452
## sd(time)                0.09      0.07     0.00     0.26 1.00      425      281
## cor(Intercept,time)     0.28      0.52    -0.80     0.98 1.01      642      424
##
## Population-Level Effects:
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept     0.10      0.22    -0.33     0.55 1.00      836      720
## Gpa_pm       -0.08      0.08    -0.23     0.08 1.00      779      673
## Gpa_pmc      -0.24      0.15    -0.53     0.05 1.00     1026      589
## time          0.10      0.08    -0.06     0.28 1.00     1242      588
## ethnic1       0.12      0.12    -0.12     0.35 1.00      981      698
##
## 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).
```

model2: two way interactions:

```
#test random slopes
# a model with only random slope of time
m11 <- brm(bf(pop ~ (Gpa_pm+ Gpa_pmc + time ) * ethnic + Gpa_pmc * time + Gpa_pm * ti
me + (time|id),sigma ~ 0 ), data=dat_long,
             seed = 1150,
             file = "final11",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc (in addition to time)
m12 <- brm(bf(pop ~ (Gpa_pm+ Gpa_pmc + time ) * ethnic + Gpa_pmc * time + Gpa_pm * ti
me + (Gpa_pmc + time|id),sigma ~ 0 ), data=dat_long,
             seed = 1151,
             file = "final12",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc * time
m13 <- brm(bf(pop ~ (Gpa_pm+ Gpa_pmc + time ) * ethnic + Gpa_pmc * time + Gpa_pm * ti
me + (Gpa_pmc * time|id),sigma ~ 0 ), data=dat_long,
             seed = 1151,
             file = "final13",
             chains = 2L, iter = 1000L)



loo(m11, m12, m13)
```

```
## Output of model 'm11':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -894.8 19.7
## p_loo        85.5  7.0
## looic      1789.5 39.5
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     629   94.9%   156
##  (0.5, 0.7]   (ok)        31    4.7%   62
##    (0.7, 1]   (bad)        3    0.5%   48
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm12':
```

```
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -895.4 19.7
## p_loo        87.0  7.3
## looic      1790.7 39.5
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     616   92.9%   92
##  (0.5, 0.7]   (ok)        44    6.6%   72
##    (0.7, 1]   (bad)        3    0.5%   31
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm13':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -894.5 19.5
## p_loo        87.2  7.2
## looic      1789.0 39.0
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     624   94.1%   70
##  (0.5, 0.7]   (ok)        36    5.4%   53
##    (0.7, 1]   (bad)        3    0.5%   46
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##     elpd_diff se_diff
## m13  0.0       0.0
## m11 -0.2       0.8
## m12 -0.8       0.8
```

```
# m13 has smaller looic value.
# only include random slope of time * Gpa_pmc  (opt to m13)


#model results
summary(m13)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = log
## Formula: pop ~ (Gpa_pm + Gpa_pmc + time) * ethnic + Gpa_pmc * time + Gpa_pm * time
+ (Gpa_pmc * time | id)
##          sigma ~ 0
##    Data: dat_long (Number of observations: 663)
##   Draws: 2 chains, each with iter = 1000; warmup = 500; thin = 1;
##          total post-warmup draws = 1000
##
## Group-Level Effects:
## ~id (Number of levels: 335)
##                             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)                   0.52      0.07     0.39     0.64 1.00      411
## sd(Gpa_pmc)                     0.22      0.18     0.01     0.70 1.01      269
## sd(time)                        0.09      0.07     0.00     0.26 1.00      357
## sd(Gpa_pmc:time)                0.47      0.39     0.02     1.51 1.02      202
## cor(Intercept,Gpa_pmc)         -0.14      0.44    -0.86     0.73 1.00      891
## cor(Intercept,time)             0.19      0.43    -0.74     0.87 1.00      827
## cor(Gpa_pmc,time)              -0.07      0.45    -0.82     0.76 1.00      595
## cor(Intercept,Gpa_pmc:time)     0.00      0.41    -0.77     0.77 1.00      656
## cor(Gpa_pmc,Gpa_pmc:time)      -0.23      0.46    -0.92     0.71 1.01      272
## cor(time,Gpa_pmc:time)         -0.07      0.46    -0.84     0.77 1.00      380
##                             Tail_ESS
## sd(Intercept)                    365
## sd(Gpa_pmc)                      497
## sd(time)                         552
## sd(Gpa_pmc:time)                 293
## cor(Intercept,Gpa_pmc)           484
## cor(Intercept,time)              638
## cor(Gpa_pmc,time)                827
## cor(Intercept,Gpa_pmc:time)      576
## cor(Gpa_pmc,Gpa_pmc:time)        707
## cor(time,Gpa_pmc:time)           643
##
## Population-Level Effects:
##                 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept           0.44      0.33    -0.21     1.09 1.00      575      848
## Gpa_pm             -0.18      0.12    -0.41     0.06 1.00      565      818
## Gpa_pmc            -0.26      0.30    -0.81     0.33 1.00      439      532
```

```
## time                -0.34      0.35    -1.01     0.29 1.00      564      793
## ethnic1             -0.26      0.48    -1.19     0.74 1.00      486      600
## Gpa_pm:ethnic1       0.10      0.15    -0.23     0.39 1.01      501      613
## Gpa_pmc:ethnic1      0.14      0.30    -0.49     0.72 1.01      627      749
## time:ethnic1         0.19      0.22    -0.23     0.63 1.00      561      682
## Gpa_pmc:time        -0.03      0.39    -0.79     0.70 1.00      463      769
## Gpa_pm:time          0.10      0.13    -0.14     0.35 1.00      421      745
##
## 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).
```

model3: 3 way interaction

```
#test random slopes
# a model with only random slope of time
m14 <- brm(bf(pop ~ (Gpa_pm+ Gpa_pmc) * time * ethnic + (time|id),sigma ~ 0 ), data=d
at_long,
             seed = 1150,
             file = "final14",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc (in addition to time)
m15 <- brm(bf(pop ~ (Gpa_pm+ Gpa_pmc) * time * ethnic + (Gpa_pmc + time|id),sigma ~ 0
), data=dat_long,
             seed = 1151,
             file = "final15",
             chains = 2L, iter = 1000L)

# test the random slope for GPA_pmc * time
m16 <- brm(bf(pop ~ (Gpa_pm+ Gpa_pmc) * time * ethnic + (Gpa_pmc * time|id),sigma ~ 0
), data=dat_long,
             seed = 1151,
             file = "final16",
             chains = 2L, iter = 1000L)



loo(m14, m15, m16)
```

```
## Output of model 'm14':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
```

```
## elpd_loo   -896.9 19.9
## p_loo        87.2  7.2
## looic      1793.7 39.7
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     617   93.1%   75
##  (0.5, 0.7]   (ok)        42    6.3%   50
##    (0.7, 1]   (bad)        4    0.6%   25
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm15':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -896.8 19.8
## p_loo        89.1  7.5
## looic      1793.5 39.7
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     616   92.9%   128
##  (0.5, 0.7]   (ok)        44    6.6%   52
##    (0.7, 1]   (bad)        2    0.3%   47
##    (1, Inf)   (very bad)   1    0.2%   19
## See help('pareto-k-diagnostic') for details.
##
## Output of model 'm16':
##
## Computed from 1000 by 663 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -898.1 20.0
## p_loo        89.7  7.6
## looic      1796.3 39.9
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     618   93.2%   148
```

```
##  (0.5, 0.7]   (ok)        41    6.2%   40
##    (0.7, 1]   (bad)        4    0.6%   72
##    (1, Inf)   (very bad)   0    0.0%   <NA>
## See help('pareto-k-diagnostic') for details.
##
## Model comparisons:
##     elpd_diff se_diff
## m15  0.0       0.0
## m14 -0.1       0.9
## m16 -1.4       1.1
```

```
# m15 has smaller looic value.
#  include random slope of time and GPA_pmc (opt to m15)


#model results
summary(m15)
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = log
## Formula: pop ~ (Gpa_pm + Gpa_pmc) * time * ethnic + (Gpa_pmc + time | id)
##          sigma ~ 0
##    Data: dat_long (Number of observations: 663)
##   Draws: 2 chains, each with iter = 1000; warmup = 500; thin = 1;
##          total post-warmup draws = 1000
##
## Group-Level Effects:
## ~id (Number of levels: 335)
##                         Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)               0.52      0.07     0.39     0.65 1.00      502
## sd(Gpa_pmc)                 0.15      0.12     0.01     0.44 1.00      847
## sd(time)                    0.10      0.07     0.00     0.27 1.00      457
## cor(Intercept,Gpa_pmc)     -0.14      0.46    -0.88     0.75 1.00     1387
## cor(Intercept,time)         0.21      0.47    -0.74     0.94 1.00     1028
## cor(Gpa_pmc,time)          -0.07      0.50    -0.88     0.87 1.00      859
##                         Tail_ESS
## sd(Intercept)                586
## sd(Gpa_pmc)                  540
## sd(time)                     298
## cor(Intercept,Gpa_pmc)       729
## cor(Intercept,time)          769
## cor(Gpa_pmc,time)            610
##
## Population-Level Effects:
##                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## Intercept                0.64      0.38    -0.08     1.37 1.01      312
```

```
## Gpa_pm                   -0.26      0.15    -0.54     0.01 1.00      324
## Gpa_pmc                  -0.36      0.35    -1.09     0.30 1.00      548
## time                     -0.71      0.48    -1.64     0.19 1.01      326
## ethnic1                  -0.71      0.63    -2.00     0.49 1.01      290
## Gpa_pm:time               0.25      0.19    -0.12     0.62 1.01      288
## Gpa_pmc:time              0.16      0.54    -0.89     1.19 1.00      504
## Gpa_pm:ethnic1            0.25      0.21    -0.13     0.67 1.01      280
## Gpa_pmc:ethnic1           0.29      0.51    -0.73     1.31 1.00      564
## time:ethnic1              1.04      0.78    -0.47     2.70 1.01      278
## Gpa_pm:time:ethnic1      -0.29      0.26    -0.81     0.21 1.01      240
## Gpa_pmc:time:ethnic1     -0.35      0.80    -1.93     1.28 1.00      576
##                       Tail_ESS
## Intercept                  506
## Gpa_pm                     580
## Gpa_pmc                    555
## time                       570
## ethnic1                    545
## Gpa_pm:time                490
## Gpa_pmc:time               597
## Gpa_pm:ethnic1             536
## Gpa_pmc:ethnic1            680
## time:ethnic1               646
## Gpa_pm:time:ethnic1        504
## Gpa_pmc:time:ethnic1       682
##
## 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).
```