Homework 8

Psych 5068

Emorie Beck

April 14, 2018

Contents

Workspace

Packages

Data

```
source("https://raw.githubusercontent.com/emoriebeck/homeworks/master/table_fun.R")
data_url <- "https://raw.githubusercontent.com/emoriebeck/homeworks/master/homework8/iq(2).csv"
dat      <- data_url %>% read.csv %>% tbl_df
```

Question 1

First, create a grand mean centered version of the family size variable. Name it fam_size_GMC.

```
dat <- dat %>%
  mutate(fam_size_GMC = fam_size - mean(fam_size, na.rm = T))
```

Question 2

2. Begin wiht the following model (Model 1), which treats all coefficients as random. Is there any evidence for birth order or family size effects in this analysis?

Table 1: Question 2: Model 1

	Score	
Term	b	CI
Fixed Parts		
verbal	30.31	[29.39, 31.25]
spatial	30.27	[29.43, 30.95]
verbal:order	-0.93	[-2.39, -0.37]
spatial:order	-0.38	[-1.01, 0.06]
verbal:fam_size_GMC	0.01	[-0.39, 0.53]
spatial:fam_size_GMC	-0.06	[-0.73, 0.12]
verbal:order:fam_size_GMC	-0.00	[-0.49, 0.27]
spatial:order:fam_size_GMC	0.14	[-0.17, 0.66]
Random Parts		
τ_{00}	9.05	[6.19, 9.56]
τ_{11}	8.68	[7.65, 9.21]
R^2_m	0.00	
R^2_c	0.72	

$$score = \pi_{0jk} * verbal + \pi_{1jk} * spatial$$

$$pi_{0jk} = \beta_{00k} + \beta_{01k} * order + r_{00k}$$
$$pi_{1jk} = \beta_{10k} + \beta_{11k} * order + r_{10k}$$

$$\beta_{00k} = \gamma_{000} + \gamma_{001} * famsize + u_{00k}$$
$$\beta_{01k} = \gamma_{010} + \gamma_{011} * famsize + u_{01k}$$
$$\beta_{10k} = \gamma_{100} + \gamma_{101} * famsize + u_{10k}$$
$$\beta_{11k} = \gamma_{110} + \gamma_{111} * famsize + u_{11k}$$

```
Model_1 <- lmer(score ~
                -1 + verbal + spatial + # level 1
                order:verbal + order:spatial +  # level 2
                fam_size_GMC:verbal + fam_size_GMC:spatial + # level 3
                fam_size_GMC:order:verbal + fam_size_GMC:order:spatial + # level 3
                (-1 + verbal + spatial|child_unique) +
                (-1 + verbal + spatial + order:verbal + order:spatial|family),
                data=dat)
tab1    <- table_fun(Model_1)

tab1 %>% select(-type) %>%
  # mutate(term = str_replace_all(term, "\\_", "\\\\_")) %>%
  kable(., "latex", booktabs = T, escape = F,
        col.names = c("Term", "b", "CI"),
        caption = "Question 2: Model 1") %>%
  kable_styling(full_width = F) %>%
  column_spec(2:3, width = "2cm") %>%
  group_rows("Fixed Parts",1,8) %>%
  group_rows("Random Parts",9,10) %>%
  group_rows("Model Terms",11,12) %>%
  add_header_above(c(" " = 1, "Score" = 2))
```

Based on the model, there is no evidence for an effect of family size.

Table 2: Question 3: Model 2

	Score	
Term	b	CI
Fixed Parts		
verbal	30.30	[29.22, 30.74]
spatial	30.27	[29.12, 30.60]
verbal:order	-0.93	[-1.65, -0.19]
spatial:order	-0.40	[-0.72, 0.70]
Random Parts		
τ_{00}	9.05	[7.77, 10.14]
τ_{11}	8.66	[6.84, 9.43]
R^2_m	0.00	
R^2_c	0.72	

Question 3

Modify the first model (call the modification Model 2) in a way that will provide an omnibus test for family size effect when the two models are compared. (Hint: The models should differ by 4 degrees of freedom). What does this model comparison tell you about the presence of family size effects in this sample?

```
Model_2 <- lmer(score ~
                -1 + verbal + spatial +         # level 1
                order:verbal + order:spatial +  # level 2
                (-1 + verbal + spatial|child_unique) +
                (-1 + verbal + spatial + order:verbal + order:spatial|family),
                data=dat)
tab2    <- table_fun(Model_2)

tab2 %>% select(-type) %>%
  # mutate(term = str_replace_all(term, "\\_", "\\\\_")) %>%
  kable(., "latex", booktabs = T, escape = F,
        col.names = c("Term", "b", "CI"),
        caption = "Question 3: Model 2") %>%
  kable_styling(full_width = F) %>%
  column_spec(2:3, width = "2cm") %>%
  group_rows("Fixed Parts",1,4) %>%
  group_rows("Random Parts",5,6) %>%
  group_rows("Model Terms",7,8) %>%
  add_header_above(c(" " = 1, "Score" = 2))
```

```
anova(Model_1, Model_2)
```

```
## Data: dat
## Models:
## Model_2: score ~ -1 + verbal + spatial + order:verbal + order:spatial +
## Model_2:     (-1 + verbal + spatial | child_unique) + (-1 + verbal + spatial +
## Model_2:     order:verbal + order:spatial | family)
## Model_1: score ~ -1 + verbal + spatial + order:verbal + order:spatial +
## Model_1:     fam_size_GMC:verbal + fam_size_GMC:spatial + fam_size_GMC:order:verbal +
## Model_1:     fam_size_GMC:order:spatial + (-1 + verbal + spatial | child_unique) +
## Model_1:     (-1 + verbal + spatial + order:verbal + order:spatial | family)
```

Table 3: Question 4: Model 3

	Score	
Term	b	CI
Fixed Parts		
verbal	29.84	[29.19, 31.30]
spatial	30.07	[29.56, 31.24]
Random Parts		
τ_{00}	9.22	[8.10, 9.79]
τ_{11}	8.71	[7.66, 9.34]
R^2_m	0.00	
R^2_c	0.72	

```
##         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## Model_2 18 12976 13080 -6470.0    12940
## Model_1 22 12984 13111 -6469.8    12940 0.3025      4     0.9897
```

There appear to be no family effects in this sample.

Question 4

Now modify Model 2 to remove all birth order effects (call the new model Model 3) and conduct a model comparison to Model 2.

```
Model_3 <- lmer(score ~
                -1 + verbal + spatial + # level 1
                (-1 + verbal + spatial | child_unique) +
                (-1 + verbal + spatial | family),
                data=dat)
tab3    <- table_fun(Model_3)

tab3 %>% select(-type) %>%
  # mutate(term = str_replace_all(term, "\\_", "\\\\_")) %>%
  kable(., "latex", booktabs = T, escape = F,
        col.names = c("Term", "b", "CI"),
        caption = "Question 4: Model 3") %>%
  kable_styling(full_width = F) %>%
  column_spec(2:3, width = "2cm") %>%
  group_rows("Fixed Parts",1,2) %>%
  group_rows("Random Parts",3,4) %>%
  group_rows("Model Terms",5,6) %>%
  add_header_above(c(" " = 1, "Score" = 2))
```

```
anova(Model_3, Model_2)
```

```
## Data: dat
## Models:
## Model_3: score ~ -1 + verbal + spatial + (-1 + verbal + spatial | child_unique) +
## Model_3:     (-1 + verbal + spatial | family)
## Model_2: score ~ -1 + verbal + spatial + order:verbal + order:spatial +
## Model_2:     (-1 + verbal + spatial | child_unique) + (-1 + verbal + spatial +
## Model_2:     order:verbal + order:spatial | family)
```

```
##         Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## Model_3  9 12964 13016 -6472.9    12946
## Model_2 18 12976 13080 -6470.0    12940   5.9      9     0.7499
```

Part A

What can you conclude from this comparison? There do not appear to be any birth order effects.

Part B

Compare Model 3 to Model 1. What is being tested and what does it offer beyond the previous two model comparisons?

```
anova(Model_3, Model_1)
```

```
## Data: dat
## Models:
## Model_3: score ~ -1 + verbal + spatial + (-1 + verbal + spatial | child_unique) +
## Model_3:     (-1 + verbal + spatial | family)
## Model_1: score ~ -1 + verbal + spatial + order:verbal + order:spatial +
## Model_1:     fam_size_GMC:verbal + fam_size_GMC:spatial + fam_size_GMC:order:verbal +
## Model_1:     fam_size_GMC:order:spatial + (-1 + verbal + spatial | child_unique) +
## Model_1:     (-1 + verbal + spatial + order:verbal + order:spatial | family)
##         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## Model_3  9 12964 13016 -6472.9    12946
## Model_1 22 12984 13111 -6469.8    12940 6.2025     13     0.9385
```

This is testing whether a model that only estimates mean verbal and spatial scores, as well as unique means for each family and child within a family is a better model than one that adjusts the estimated means based on birth order and the number of children in the family. Neither of these seem to impact mean verbal or spatial ability.

Question 5

Using the most parsimonious of the three previous models, modify it (call it Model 4) to provide a model comparison that tests the significance of the true score correlation between verbal and spatial abilities at the level of the child.

```
Model_4 <- lmer(score ~
                -1 + verbal + spatial + # level 1
                (-1 + verbal + spatial || child_unique) +
                (-1 + verbal + spatial | family),
                data=dat)

Model_5 <- lmer(score ~
                -1 + verbal + spatial + # level 1
                (1 | child_unique) +
                (1 | family),
                data=dat)
tab4    <- table_fun(Model_4)
tab4 %>% select(-type) %>%
  # mutate(term = str_replace_all(term, "\\_", "\\\\_")) %>%
```

Table 4: Question 4: Model 3

	Score	
Term	b	CI
Fixed Parts		
verbal	29.83	[29.18, 30.58]
spatial	30.07	[29.72, 30.91]
Random Parts		
τ_{00}	9.21	[7.60, 9.72]
R^2_m	0.00	
R^2_c	0.72	

```
  kable(., "latex", booktabs = T, escape = F,
        col.names = c("Term", "b", "CI"),
        caption = "Question 4: Model 3") %>%
  kable_styling(full_width = F) %>%
  column_spec(2:3, width = "2cm") %>%
  group_rows("Fixed Parts",1,2) %>%
  group_rows("Random Parts",3,3) %>%
  group_rows("Model Terms",4,5) %>%
  add_header_above(c(" " = 1, "Score" = 2))
```

```
t_cor <- (sigma(Model_5)^2 - sigma(Model_4)^2)/sigma(Model_5)^2
```

```
anova(Model_4, Model_5)
```

```
## Data: dat
## Models:
## Model_5: score ~ -1 + verbal + spatial + (1 | child_unique) + (1 | family)
## Model_4: score ~ -1 + verbal + spatial + ((0 + verbal | child_unique) +
## Model_4:     (0 + spatial | child_unique)) + (-1 + verbal + spatial |
## Model_4:     family)
##         Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## Model_5  5 13264 13294 -6627.3    13254
## Model_4  8 12996 13042 -6490.0    12980 274.6      3  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The true score correlation is 0.33.