

PSYC 5670: Homework 3

J DeVaney

Question 1:

```
#library chunk not included
setwd("~/Desktop/Files/HLM/datasets/")
ext <- read_sas("external_pp.sas7bdat")

#converting ID to factor
ext$ID <- as.factor(ext$ID)
```

Question 2:

```
#descriptives table for EXTERNAL
ext %>%
  group_by(TIME) %>%
  summarise(n = n(), mean = mean(EXTERNAL), sd = sd(EXTERNAL), min = min(EXTERNAL), max
= max(EXTERNAL)) %>%
  round(digits = 2) %>%
  kable()
```

TIME	n	mean	sd	min	max
0	45	13.87	12.89	0	50
1	45	13.20	12.17	0	57
2	45	12.18	12.62	0	51
3	45	13.24	13.60	0	48
4	45	11.58	11.45	0	43
5	45	13.71	13.17	0	56

Question 3:

$$L1: EXTERNAL_{ij} = B_{0j} + r_{ij}$$

$$L2: B_{0j} = \text{Gamma}_{00} + \text{TIME}_j$$

$$\text{Reduced: } EXTERNAL_{ij} = \text{Gamma}_{00} + \text{TIME}_j + r_{ij}$$

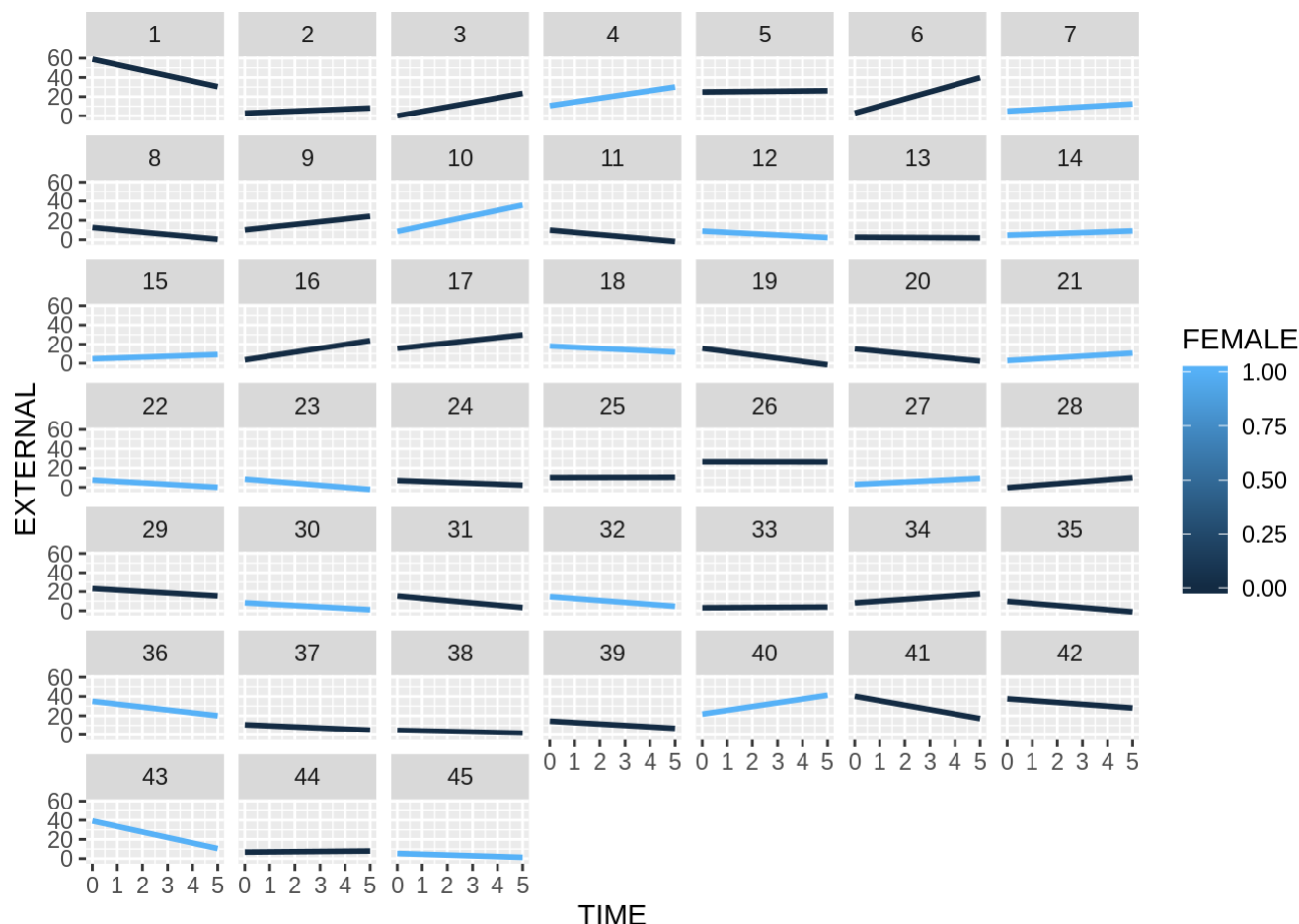
Gamma_{00} : marginal mean of EXTERNAL

TIME_j : deviation of cluster mean for given timepoint from grand mean

r_{ij} : deviation of i th response from j th cluster mean

Question 4:

```
ext %>%
  ggplot(aes(x = TIME, y = EXTERNAL, color = FEMALE)) +
  geom_smooth(method = lm, se = F) +
  facet_wrap(ext$ID)
```



Facet by ID makes obvious the variation in slope and intercept for each person. The arrangement makes it difficult to interpret overall differences between male/female intercepts and slopes. A multilevel model will likely be able to more effectively account for variance, and, as seen in the plots in the next question, will reflect a difference in the intercepts and slopes between males and females.

Question 5:

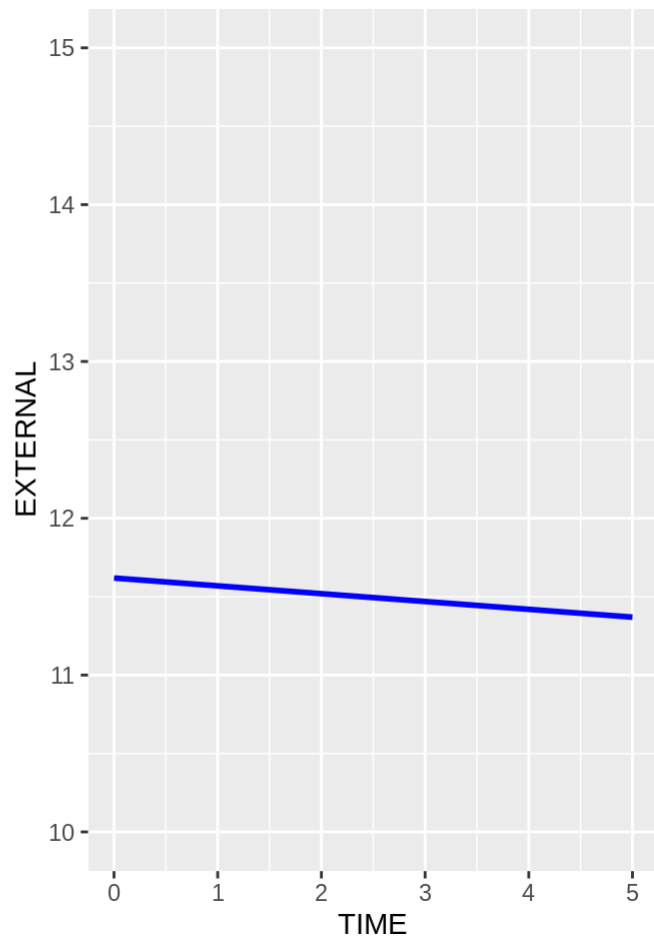
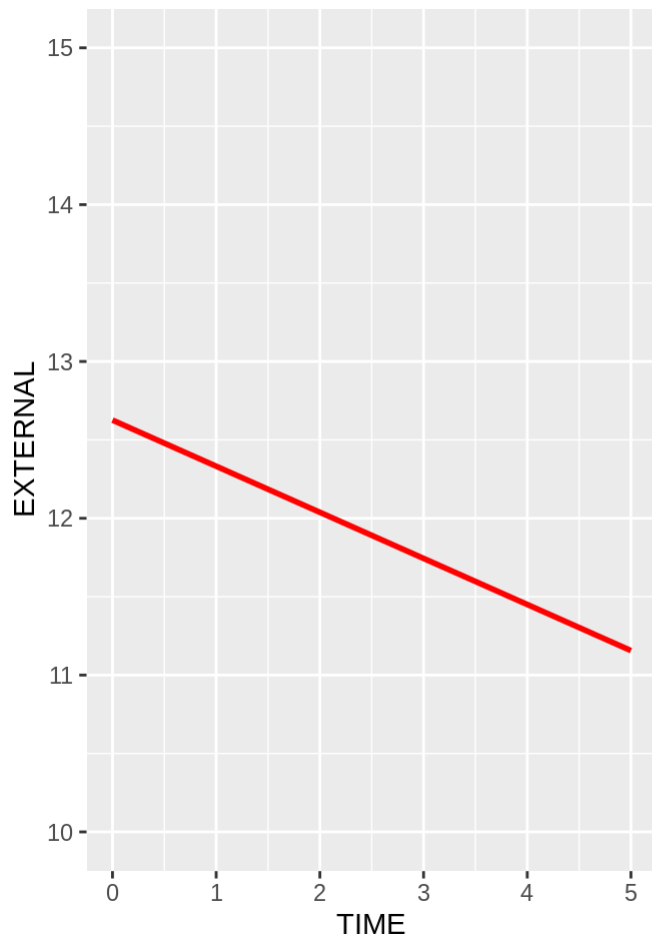
```
p1 <- filter(ext, FEMALE == 1) %>%
  ggplot(aes(x = TIME, y = EXTERNAL)) +
  geom_smooth(method = lm, se = F, color = "red") +
  ylim(10, 15)

p2 <- filter(ext, FEMALE == 0) %>%
  ggplot(aes(x = TIME, y = EXTERNAL)) +
  geom_smooth(method = lm, se = F, color = "blue") +
  ylim(10, 15)

cowplot::plot_grid(p1, p2)
```

```
## Warning: Removed 85 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 142 rows containing non-finite values (stat_smooth).
```



Question 6

```
m0 <- lmer(data = ext, EXTERNAL ~ TIME + (1 + TIME|ID),
            control = lmerControl(optimizer = "bobyqa",
                                   optCtrl = list(maxfun = 200000)))
tab_model(m0, show.aic = T, show.r2 = F, show.ci = F, show.se = T)
```

EXTERNAL			
Predictors	Estimates	std. Error	p
(Intercept)	13.29	1.86	<0.001
TIME	-0.13	0.42	0.755

Random Effects

σ^2	53.72
τ_{00} ID	126.97
τ_{11} ID.TIME	4.87
ρ_{01} ID	-0.52
ICC	0.66

N _{ID}	45
Observations	270
AIC	2001.015

Question 7

```
m1 <- lmer(data = ext, EXTERNAL ~ 1 + FEMALE + (1 + TIME|ID),
           control = lmerControl(optimizer = "bobyqa",
                                optCtrl = list(maxfun = 200000)))
```

Question 8

```
m2 <- lmer(data = ext, EXTERNAL ~ 1 + TIME*FEMALE + (1 + TIME|ID),
           control = lmerControl(optimizer = "bobyqa",
                                optCtrl = list(maxfun = 200000)))
tab_model(m1, m2, show.aic = T, show.r2 = F, show.ci = F, show.se = T)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates the c
onditional random effect variances.
## Solution: Respecify fixed structure!
```

	EXTERNAL			EXTERNAL		
Predictors	Estimate	std. Error	p	Estimate	std. Error	p
(Intercept)	13.44	1.92	<0.001	13.99	2.37	<0.001
FEMALE	-1.29	3.12	0.680	-1.86	3.86	0.631
TIME				-0.21	0.54	0.692
TIME:FEMALE				0.22	0.88	0.803
Random Effects						
σ^2	53.72			53.72		
τ_{00}	127.49 _{ID}			129.73 _{ID}		
τ_{11}	4.71 _{ID.TIME}			5.04 _{ID.TIME}		
ρ_{01}	-0.51 _{ID}			-0.52 _{ID}		
ICC	0.70			0.67		
N	45 _{ID}			45 _{ID}		
Observations	270			270		
AIC	1996.927			1999.112		

Question 9

```
options(scipen = 99999)
mEB <- lmer(data = ext, EXTERNAL ~ 1 + (1|ID), na.action = na.omit)

#from viewing summary for list indices
EB_int <- ranef(mEB)$ID + summary(mEB)[[10]][1]

contrasts(ext$ID) <- contr.sum
```

```
lm(data = ext, EXTERNAL ~ ID) %>% summary()
#lengthy output hidden
```

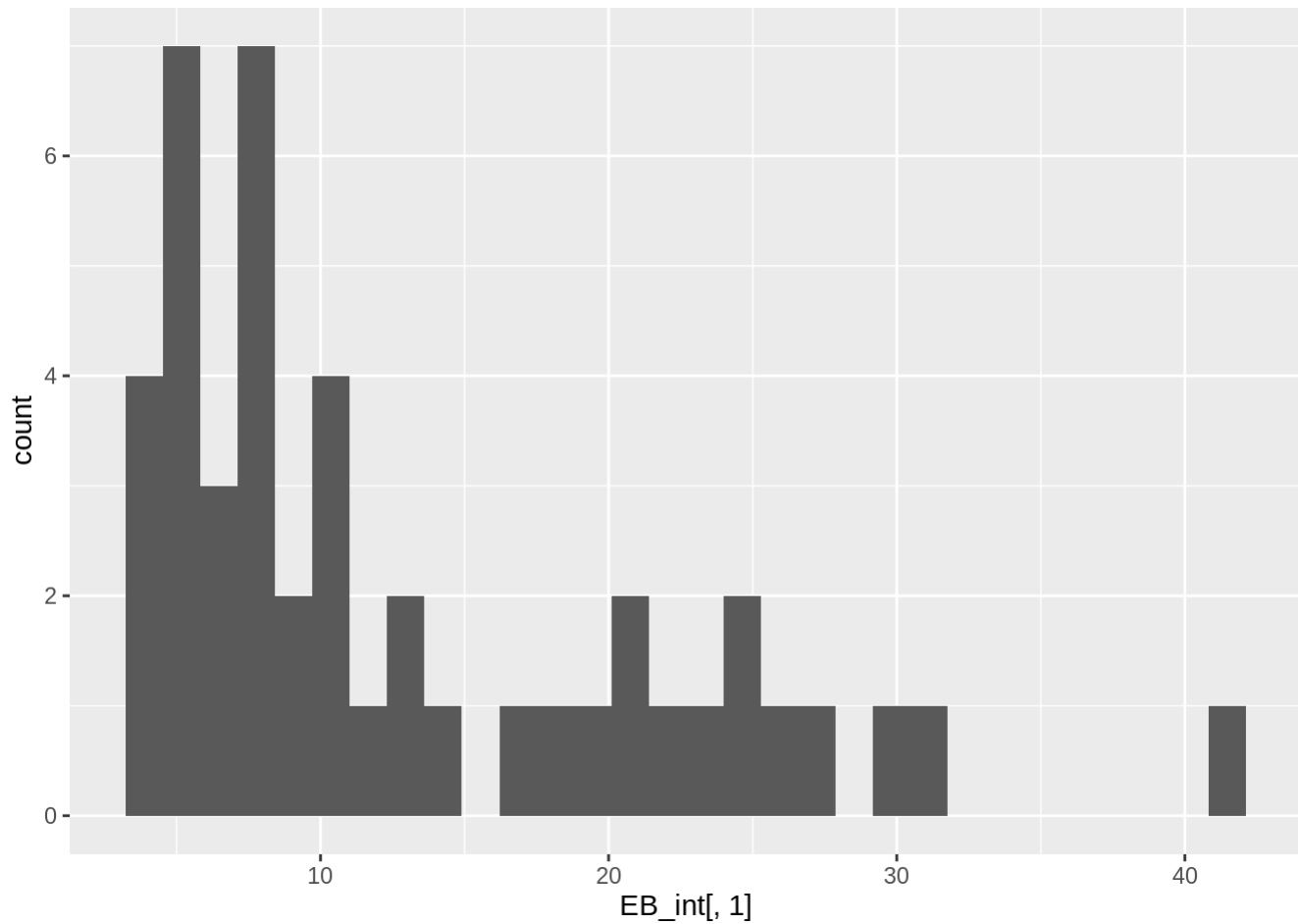
Question 10

```
# p3 <- EBdf %>%
#   ggplot(aes(x = EBdf[,1], y = EBdf[,2])) +
#   geom_point()
```

Question 11

```
hist1 <- ggplot(data = EB_int, aes(x = EB_int[,1])) +
  geom_histogram()
hist1
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
# hist2 <- ggplot(data = EB_int, aes(x = EB_int[,2])) +  
#   geom_histogram()
```

The first histogram of the intercepts shows a positive skew, giving reason to be concerned about violating an assumption of normally distributed random effects.