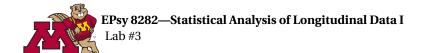# ANALYZING BASIC MODELS WITH lm() AND lmer()

This lab is intended to give you experience in using the `lm()` function for traditional linear regression (LR) and the `lmer()` function for linear mixed effect regression (LMER). The long format data previously considered will be used in both the LR and LMER analyses. This is not appropriate for LR but we use it in anticipation of moving to LMER.

You will use the `sleeplab.Rdata` data set you created in the first lab. Make a note of where this file resides on your hard drive. An Rdata file is loaded into R using the `load()` function. If you do not have this file, you must remake by following the steps in Lab 1. See Lab 1 for a complete description of the data set. The response variable is `Reaction`, which is the average reaction time for a number of cognitive tests. The time metric is `Days` $(0, 1, \ldots, 9)$, and the static predictors are `gpa` and two versions of sex, `female` $(0 = \text{male}, 1 = \text{female})$ and `female.c` ("male", "female").

Read in the Data: To load the Rdata file, type the following syntax in your script file and run it.

```
> # The following line must be tailored to your system:
> load(file = "/.../sleeplab.Rdata")
> ls()
```

/.../ indicates you replace … with the location of the file. For example,, on my computer the entire syntax would be /Users/zief0002/Documents/EPsy8282/sleeplab.Rdata. For Windows users, replace the pathname with C:\\...\\. Be sure to use double backslashes.

An indication you successfully loaded the file is the absence of any error messages. That is, success is indicated by a new prompt line in R (i.e., >). If you did not successfully read in the file, then check your syntax and try again.

The ls() function shows the objects in the Rdata file, which are data frames. If you followed the directions in Lab 1, the long format data frame is named sleep.long3. If you renamed this data frame, then be sure to use the new name in the syntax below. It is a good idea to run head(sleep.long3) to see the top of the data set.

Assignment Guidelines: In each section you are directed to produce output, which you should copy from R and paste into your word/document processor. Please label the sections as indicated below and use the question numbering as indicated. All questions are worth 1 point. There is a total of 39 points possible. There is no external R script for this lab. The R commands you will need are embedded in the lab itself.
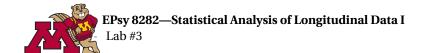
Plotting Individual Growth Curves

Though the last lab dealt extensively with graphing, it is important to always start with a look at the data. Run the syntax below to create a graph of the superimposed growth curves.

```
> library(ggplot2)
> theme_set(theme_bw())
> ggplot(data=sleep.long3, aes(x = Days, y = Reaction, groups = SubNum)) +
        geom_line() +
        theme_bw() +
        opts(title = "Individual Curves")
```

Include the graph in your write-up and answer the following questions based on the graph you produced above. You can label this as "Section: Graph I" in your write-up. Please keep the answers short.

1. Based on the graph, is there an indication you need a random effect for the intercept in a LMER model? Be specific.

2. Based on the graph, is there an indication you need a random effect for the slope in a LMER model? Be specific.

Getting Your Feet Wet: Linear Regression

Before considering the LMER model, you will work with the LR model also known as the linear model (LM). In this lab you will consider the simple case of using days of the study to predict reaction time scores. Then you will consider the case of sex as a predictor in addition to days. The predictors will first be included as single terms and then their interaction will be included.
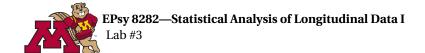
Single Quantitative Predictor

Run the syntax below, which models `Reaction` as a function of `Days`.

```
> LM.1 <- lm(Reaction ~ 1 + Days, data = sleep.long3)
> summary(LM.1)
```

Answer the following questions ~~based on the two graphs you produced immediately above.~~ You can label this as "Section: Linear Regression I" in your write-up. Please keep the answers short.

3. Using symbolic notation, including Roman letters or variable names where appropriate, write the LR model corresponding to the above syntax (see the lecture notes).

4. Provide a detailed interpretation of the estimated intercept.

5. Provide a detailed interpretation of the estimated slope.

6. What does the p-value for the slope indicate in this case? Be specific.

7. What assumptions are made in computing the p-values in this case? Be specific.

ANCOVA Model

In the second model you will use sex and study days as predictors of reaction time. In LR this is sometimes referred to as an analysis of covariance (ANCOVA) because one predictor is categorical and another is quantitative. In R, it is better to save a categorical variable as a factor variable, and then use the factor variable in the syntax. Before estimating the LR model, save `female.c` as the factor variable `female.f` in the data frame.
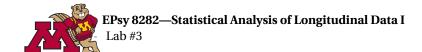
```
> sleep.long3$female.f <- factor(sleep.long3$female.c, labels = c("Male", "Female"))
> head(sleep.long3)
```

Having saved the factor variable, run the following syntax for the ANCOVA model.

```
> LM.2 <- lm(Reaction ~ 1 + Days + female.f, data = sleep.long3)
> summary(LM2.out)
```

Answer the following questions. You can label this as "Section: Linear Regression II" in your write-up.

8. Notice the variable label name in the third row of the `Coefficients` table. Explain why the variable label is not `female.f`. Be detailed in your answer.

9. Would the results be different if `female` was used rather than `female.f` in the syntax above? Explain.

10. Compare the multiple R^2 values (not adjusted) for `LM.2` and `LM.1`. Based on these values, is there statistical evidence that `female.f` is an "important" predictor? Explain.

11. Provide a detailed interpretation of the fixed effect estimate associated with `Days`.

12. Provide a detailed interpretation of the last fixed effect estimate.

13. Write the predicted (fitted) change curve equation for females using the estimates from the output.

14. Write the predicted (fitted) change curve equation for males using the estimates from the output.

15. Based on your answers for the last two questions, explain any differences in the change curves for males and females.

16. If you were to graph the predicted (fitted) curves for females and males in this example, would the lines cross? Explain why or why not.
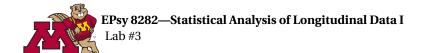
Interaction Model

The third model includes the interaction between `female.f` and `Days`, in addition to the single effects. Run the syntax below.

```
> LM.3 <- lm(Reaction ~ 1 + Days + female.f + Days:female.f, data = sleep.long3)
> summary(LM.3)
```

Answer the following questions. You can label this as "Section: Linear Regression III" in your write-up.

17. What does the following syntax do? `Days:female.f`. Be specific.

18. Notice the variable label name in the fourth row of the `Coefficients` table. Explain why the variable label is not `Days:female.f`. Be detailed in your answer.

19. Provide a detailed interpretation of the fixed effect estimate associated with `Days`.

20. Provide a detailed interpretation of the fixed effect estimate associated with `Days:female.fFemale`.

21. Based on the information in the `Days:female.fFemale` row of the `Coefficients table`, is there statistical evidence we should adopt the interaction model of this section or the ANCOVA model of the last section? Explain.

22. Write the predicted (fitted) change curve equation for females using the estimates from the output.

23. Write the predicted (fitted) change curve equation for males using the estimates from the output.

24. Based on your answers for the last two questions, explain any differences in the change curves for males and females.

25. If you were to graph the predicted (fitted) curves for females and males in this example, would the lines cross? Explain why or why not.

Linear Mixed Effects Regression

The above three LR models have LMER analogs. Estimate each LMER model and answer the questions in each section.

LMER I
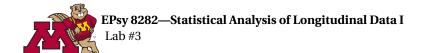
The simplest LMER model we consider has Days as the predictor. Run the following syntax.

```
> library(lme4)
> LMER.1 <- lmer(Reaction ~ Days + (1 + Days | SubNum), data = sleep.long3, REML = F)
> summary(LMER.1)
```

Answer the following questions. You can label this as "Section: LMER I" in your write-up.

26. What does the following syntax do? REML=F.

27. Using symbolic notation, including Roman letters and variable names where appropriate, write the LMER model corresponding to the above syntax (see lecture notes).

28. Consider the Random effects table in the output. Explain each of the numbers in the Variance column, i.e., interpret the values.

29. Consider the Random effects table in the output. Explain what the value in the Corr column represents and what it means in this situation (i.e., provide an interpretation for this example).

30. Consider the Random effects table in the output. Based on the estimated values, is there evidence that the second random effect belongs in the model? Explain. (Hint: no addition calculations are required.)

31. Consider the Fixed effects table in the output. Provide a detailed interpretation of the first fixed effect estimate in the (Intercept) row.

32. Consider the Fixed effects table in the output. Provide a detailed interpretation of the fixed effect estimate in the Days row.

33. Compare the fixed effects estimates of LMER.1 and LM.1. Would you come to a very different conclusion regarding the change curve based on the two sets of estimates? Explain.

LMER II

The second LMER model you will consider adds `female.f` as a static predictor. Run
the syntax below.

```
>LMER.2 <- lmer(Reaction ~ Days + female.f + (1 + Days | SubNum), data = sleep.long3, REMI
> summary(LMER.2)
```

Answer the following questions. You can label this as "Section: Random Effects" in
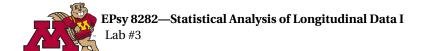your write-up.

34. Provide a detailed interpretation of the fixed effect estimate in the `female.fFemale`
   row of the `Fixed effects` table.

35. What do the *t* values in the `Fixed effects` table indicate about the relative
   importance of the predictors? Explain. (Hint: not additional calculations re-
   quired.)

36. Compare the `(Intercept)` variance in the `Random effects` table for `LMER.2`
   and `LMER.1`. Explain any difference and what accounts for the difference.

37. The `Residual` variance in the `Random effects` table did not change much
   from `LMER.1` to `LMER.2`. Explain why.

38. If you were to estimate a model that *excluded* Days and `female.f`, what would
   happen to the `Residual` variance in the `Random effects` table and why?

39. Write the predicted (fitted) change curve for females using the estimates from
   the output. (Hint: random effects are not estimated.)

40. Write the predicted (fitted) change curve for males using the estimates from
   the output. (Hint: random effects are not estimated.)

41. Based on your answers for the last two questions, explain any differences in
   the change curves for males and females.

LMER III

The last LMER model you will consider adds the interaction between Days and `female.f`.
Run the syntax below.

```
> LMER.3 <- lmer(Reaction ~ 1 + Days + female.f + Days:female.f + ( 1 + Days | SubNum),
            data = sleep.long3, REML = F)
> summary(LMER.3)
```

Answer the following questions. You can label this as "Section: LMER III" in your
write-up.

42. Compare the Days variance in the `Random effects` table for `LMER.3` and
   `LMER.2`. Explain any difference and what accounts for the difference.

43. Consider the `Days:female.fFemale` row of the `Fixed effects` table. Is there statistical evidence that the interaction term belongs in the model? Explain. (Hint: do not do any additional calculations.)

44. Write the predicted (fitted) change curve for females using the estimates from the output. (Hint: the random effects are not estimated.)

45. Write the predicted (fitted) change curve for males using the estimates from the output. (Hint: the random effects are not estimated.)

46. Based on your answers for the last two questions, explain any differences in the change curves for males and females.

Graph of Interaction

To gain experience in graphing based on `lmer()` output, you will construct a graph for malesand females based on `LMER.3`. Run the following syntax.

```
> ## Preliminaries.
> fixef(LMER.3)
> head(model.matrix(LMER.3))
> graphdat <- with(sleep.long3, data.frame(SubNum, Days, female.f,
        model.matrix(LMER.3) %*% fixef(LMER.3)))
> colnames(graphdat)[4] <- "predict"
> head(graphdat)

> ## ggplot graph.
> ggplot(data = graphdat, aes(x = Days, y = predict, line = female.f)) +
        geom_line() +
        theme_bw()
```

Include the graph in your write-up and answer the following questions. You can label this as "Section: Graph II" in your write-up.

47. Though the syntax used to obtain `predict` might be unclear, what do you think the values of `predict` represent? (Hint: see the output for `fixef()` above.)

48. Does the graph depict an interaction? Explain why or why not.