STATISTICAL ANALYSIS OF LONGITUDINAL DATA I

INFERENCE FOR STATIC PREDICTORS

LAB 4

Spring 2012

This lab focuses on inference for the static predictors, `GPA` and `Female`. As discussed in the lecture, analysis is more manageable when the shape of the change curve and the number of random effects is pre-selected. This can be done based on graphs of the data, and you will assume this has already been accomplished. That is, all models considered in this lab will have the following two features.

- A linear change curve.

- Random effects for the intercept and slope.

The models considered will have only different numbers of static predictors.

In this lab you will consider the two main methods of inference discussed in the lectures.

- Multimodel inference using information criterion.

- Step-up using null hypothesis significance testing (NHST).

To standardize the labs, many of the decisions that you would have to make as an applied researcher are made for you. The emphasis in the lab will be the interpretation of the results rather than the specific steps of the analysis.

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

```
#####################################################
## Tailor the first line 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/andrewz/Documents/EPsy8282/sleeplab.Rdata`. For Windows users, replace the pathname with `C:\\...\\`. Be sure to use double back-slashes.

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 30 points possible. There is no external R script for this lab. The R commands you will need are embedded in the lab itself.

Multimodel Approach

Suppose your research is concerned with gender differences in cognition under sleep deprivation. In addition to collecting the gender of your subjects, you also collected their grade point average (GPA) as you thought this might be a possible mediator of gender group differences. Based on theory, the extent literature, your professional expertise, and common sense, you formulate the following five working hypothesis.

1. Gender differences and GPA difference are negligible.

2. Males on average have higher initial scores than females and this difference persists over time.

3. Initial gender differences are mediated by GPA, at least partially.

4. Males on average have higher initial scores than females but the magnitude of the difference changes over time.

5. Initial gender differences and differences over time are mediated by GPA, at least partially.

You will translate these working hypotheses into LMER models and estimate the models. Then you will use the following information criteria to evaluate the models: AICc, Δ, weight of evidence (w), and evidence ratio (E). See the lecture notes for background on these measures.

The first step is to estimate the five models above with the sample data. Using the variable names from the data set, the lmer() syntax for running the five models appears below. Estimate the models and move on to the next step.

```
> library(lme4)
> Model.1 <- lmer(Reaction ~ Days + (Days | SubNum),
    data = sleep.long3,  REML = FALSE)
> Model.2 <- lmer(Reaction ~ Days + female +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
> Model.3 <- lmer(Reaction ~ Days + female + gpa +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
> Model.4 <- lmer(Reaction ~ Days * female +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
> Model.5 <- lmer(Reaction ~ Days * female +
    Days * gpa + (Days | SubNum), data = sleep.long3,
    REML = FALSE)
```

Now that the models have all been estimated, you are ready to compute the information criteria. Run the following syntax.

```
> library(bbmle)
> myIC <- ICtab(logLik(Model.1), logLik(Model.2),
    logLik(Model.3), logLik(Model.4), logLik(Model.5),
    type = "AICc", delta = TRUE, weights = TRUE,
    nobs = nrow(na.omit(sleep.long3)), sort = FALSE)
> myIC
> myIC$rweight <- max(myIC$weight)/myIC$weight
> myIC
```

It is convenient to create bar graphs of the weight of evidence and evidence ratio after sorting the results. Run the following syntax to accomplish these two tasks.

```
> library(ggplot2)
> mynames <- as.character(1:5)
> plotdata <- data.frame(I(mynames), myIC$weight,
    myIC$rweight)
> plotdata2 <- plotdata[order(plotdata$myIC.weight,
    decreasing = TRUE), ]

#######################################################
## Bar graph of weight of evidence
#######################################################

> ggplot(data = plotdata2,
      aes(x = mynames, y = myIC.weight)) +
    ylab("Weight") +
    xlab("Model") +
    geom_bar(fill = "grey80", colour = "black") +
    theme_bw() +
    scale_y_continuous(limits = c(0, 1))

#######################################################
## Bar graph of evidence ratio
#######################################################

> ggplot(data = plotdata2,
      aes(x = mynames, y = myIC.rweight)) +
    ylab("Weight") +
    xlab("Model") +
    geom_bar(fill = "grey80", colour = "black") +
    theme_bw() +
    geom_hline(aes(yintercept = 1), linetype = 2)
```

Include the graphs with your write-up. All other output should be included as well.

The final analysis you will perform is the parametric bootstrap for the evidence ratio. Be aware that the following syntax might take several minutes to run on your computer, so plan accordingly. R might look as if it is doing nothing during this time, i.e., it might look as if it is hung-up.

```
#####################################################
## ENSURES EVERYONE GETS THE SAME ANSWER:
## DO NOT USE IN PRACTICE
#####################################################

> set.seed(1234)

#####################################################

> bestr <- 4      # Number of best model (initial analysis)
> nrep <- 999     # Number of bootstrap replications
> mystorage <- numeric(nrep)

#####################################################

> for(r in 1:nrep){
    simDV <- simulate(Model.4)
    myfit <- ICtab(logLik(refit(Model.1, simDV[ ,1])),
        logLik(refit(Model.2, simDV[ ,1])),
        logLik(refit(Model.3, simDV[ ,1])),
        logLik(refit(Model.4, simDV[ ,1])),
        logLik(refit(Model.5, simDV[ ,1])),
        type = "AICc", nobs = nrow(na.omit(sleep.long3)),
        weights = TRUE, sort = FALSE)
    mystorage[r] <- max(myfit$weight) / myfit$weight[bestr]
    }

#####################################################
## Sort the resulting bootstrap ratios and Compute
## empirical quantiles:
#####################################################

> mystorage2 <- sort(mystorage)
> a <- c(0.1, 0.05, 0.01)  # a values
> b <- (nrep + 1) * (1 - a)  # b values
```

```
#####################################################
## Save and display quantiles
#####################################################

> myquant <- as.data.frame(mystorage2[b])
> colnames(myquant) <- "Quantile"
> rownames(myquant) <- c("90%", "95%", "99%")
> myquant
```

Be sure to check that no errors were displayed and that the quantiles are shown in the R terminal window.

Multimodel Questions

Answer the following questions based on the output you produced. You can label this as "Multimodel Questions" in your write-up. The last question requires you run an additional function in R.

1. What does REML=FALSE accomplish and why is it used here?

2. What does the logLik() function do and why is it used with ICtab()?

3. List the model numbers in order of fit starting with the best fitting and ending with the worst fitting: __, __, __, __, __.

4. Is the best fitting model the "true model" or a candidate for the "true model"? Explain.

5. Is the best fitting mode a "good" model? How would you check to see if it is "good"? Explain.

6. Is there any advantage in this case to using Δ and w rather than AICc? Explain.

7. Based on the weight of evidence, is it obvious which model is the best fitting? Explain. You might want to use the bar graph for your answer.

8. Interpret the weight of evidence for the best fitting model.

9. Interpret the evidence ratio for the worst fitting model.

10. Based on the bar graph of the evidence ratio, are there any models you would consider as *definitely* rejected? Explain why or why not.

11. Based on the bar graph of the evidence ratio, are there any models you would consider as *definitely* retained? Explain.

12. Based on the parametric bootstrap results, which models would you retain and which models would you reject based on the 0.95 quantile.

13. In terms of information theory, what does it mean to reject a model in the multimodel approach?

14. In a replication of this study, would the best fitting model again be the best fitting? Explain.

5

15. Use `summary()` to print the fitted output object for the best fitting model. Using the fixed effects estimates, what is the exact nature of the differences referred to in the verbal description of the model? (Use the estimates to clarify the nature of the effects.)

Step-Up using NHST

As discussed in this course, the multimodel approach has certain advantages, such as the direct evaluation of predata models. A less attractive but popular alternative to the multimodel approach is a NHST strategy. In this lab you will consider a step-up planned approach using NHST.

Suppose that based on theory, the extent literature, your professional expertise, and common sense, you plan to carry out statistical tests in the following steps. You will begin with a model that has no static predictors (`Days` is the only predictor). Then you will perform statistical tests for the following four steps:

- Add `female` as a single effect (intercept effect).

- Add `gpa` as a single effect.

- Add `female:Days` interaction (slope effect).

- Add `gpa:Days` interaction.

The predictor terms will be accumulated regardless their significance, though you are to make judgments of the results in the questions that follow.

The models for the steps will be numbered from 1 to 5, with the first model being the one excluding all static predictors. The syntax below estimates the models that are required for the steps. Run the syntax and then move on to the testing.

```
> NHST.1 <- lmer(Reaction ~ Days + (Days | SubNum),
    data = sleep.long3, REML = FALSE)
> NHST.2 <- lmer(Reaction ~ Days + female +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
> NHST.3 <- lmer(Reaction ~ Days + female + gpa +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
> NHST.4 <- lmer(Reaction ~ Days * female + gpa +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
> NHST.5 <- lmer(Reaction ~ Days * female + Days * gpa +
    (Days | SubNum), data = sleep.long3, REML = FALSE)
```

Now you will use `anova()` to carry out the analysis steps as described above. Run the following syntax.

```
> an.1 <- anova(NHST.1, NHST.2)
> an.2 <- anova(NHST.2, NHST.3)
> an.3 <- anova(NHST.3, NHST.4)
> an.4 <- anova(NHST.4, NHST.5)
```

It is convenient to put the results in a table. The following syntax accomplishes this.

```
> mytable <- rbind(an.1[2, 5:7], an.2[2, 5:7], an.3[2, 5:7],
    an.4[2, 5:7])
> rownames(mytable) <- c("Step 1", "Step 2", "Step 3", "Step
    4")
> mytable
```

Be sure to include the output in your write-up.

NHST Questions

Answer the following questions based on the output you produced. You can label this as "NHST Questions" in your write-up. **Your answers should reflect the fact that NHST was used for the analysis. Where applicable, consider a result statistically significant when $p \leq 0.05$.**

16. What does the `Df` column of `mytable` list, and why are all the values 1?

17. In words, what is the null hypothesis in step 1?

18. In words, what is the null hypothesis in step 4?

19. Is there evidence that you should include a gender intercept effect in the model? Explain.

20. Is there evidence that gpa mediates an intercept effect of gender? Explain.

21. Is there evidence that you should include a gender interaction effect in the model? Explain.

22. Is there evidence that you should include a gpa interaction effect in the model? Explain.

23. Based on all the steps, which one of the models (1-5) *in Section* **??** would you select and why?

24. From a *multimodel approach* perspective, is there any problem in selecting only a single model in the analysis? Explain.

25. From a *multimodel approach* perspective, is there any problem in using `NHST.1` in the step-up analysis? Explain.

26. Are models such as `NHST.1` required in other NHST approaches? Explain.

27. In the step-up approach you ran several statistical tests. Is this an issue from a NHST perspective? Is multiple testing an issue with the multimodel approach? Explain.

28. Step 2 of your analysis yielded the smallest observed chi-squared value. Does this mean that the model with female and gpa intercept effects is the worst fitting model? Explain.

29. There is a practical reason why you were not asked to run the parametric bootstrap in the step-up procedure. What do you think it is? What is the advantage in this regard of using the parametric bootstrap with the multimodel approach?

30. In step 2 of the analysis, do you think the parametric bootstrap would lead to a different conclusion about the null hypothesis? Explain.