Interaction Models II

EPsy 8251

Assignment #11

You will use the data set *FCI-2014.csv* to examine whether a set of predictors predicts variation in FCI. In this assignment, the natural logarithm of the variable `fci` will be used as the outcome. Please submit your responses to each of the questions below in a printed document. All graphics should be resized so that they do not take up more room than necessary and all should have an appropriate caption. Any equations should be appropriately typeset within the document. There are 12 points possible for the assignment (each question is worth one point).

PREPARING THE DATA

There are no points for anything in this section. However, this preparation is necessary to complete the remainder of the assignment.

- Compute the natural logarithm of FCI as a new variable (i.e., another column in the data). Name it `Lfci`.
- Compute the current age of each stadium as a new variable (i.e., another column in the data). Name it `ageStadium`.

FITTING MODELS

Fit the following regression models using R. There are no points for anything in this section. You will use the output from the fitted models to answer the questions in the assignment.

Model 1: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL`

Model 2: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL + league`

Model 3: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL + league + ageStadium:league`

Model 4: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL + league + ageStadium:league + ageStadium`^{2}`:league`

Model 5: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL + league + ageStadium:COL`

Model 6: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL + league + ageStadium:COL + ageStadium`^{2} `:COL`

Model 7: `Lfci ~ ageStadium + ageStadium`^{2} `+ COL + league + ageStadium:league + ageStadium`^{2}`:league + ageStadium:COL + ageStadium`^{2}`:COL`

INTERACTION EFFECT BETWEEN LEAGUE AND STADIUM AGE AND COST-OF-LIVING AND STADIUM AGE

1. Use the `summary()` function to examine the output from Model 7. Which coefficient(s) are not defined because of singularities?
2. What does having a singularity suggest about the specification of our model? Explain. (Note" You may need to Google (or look up in a book) what a “singularity“ is in the world of statistical linear models to answer this question.).

3. Remove the interaction term between the quadratic effect of stadium age and league (keep all the remaining effects in the model; this will be referred to as Model 8). Re-examine the output from the `summary()` function. Are all of the coefficients now estimated in this model?

4. Should all of the predictors be retained? Explain. If not, drop the predictors you believe should be removed from the model and fit a "final" model. (This model will be referred to as Model 9. If you believe all of the predictors from Model 8 should be retained, Model 9 = Model 8.)

5. Create a plot of the fitted regression equations for Model 9. In this plot, show the effect of stadium age on the x-axis and the effect of league through different lines. Also show the effect of cost-of-living through different lines.

6. Use the plot to explain the effect of stadium age on FCI. (Be careful since there are interactions.)

SOME STUFF[1] FOR PUBLICATION

7. Present all nine (or eight) models in a regression table. This table should include the estimates of the regression coefficients and standard errors, as well as some indication of statisticail significane. It should also include quantitative sumamry information (e.g., F, R^2, etc.) for each model. Be sure your table has an appropriate caption. This table should be presented on a page that is landscape oriented (not portrait oriented). If necessary, the table can be presented on multiple pages. (If you use multiple pages, be sure to have the appropriate caption, etc. on the extra pages). **(2pts.)**

8. Carry out the analysis to examine all of the pairwise league comparisons (based on Model 9), controlling for the linear and quadratic effects of stadium age, and the effect of cost-of-living. Report the estimated differences, the unadjusted p-values for those differences, and the Benjamani–Hochberg adjusted p-values for the differences in a table. Be sure your table has an appropriate caption. **(2pts.)**

9. Write 2–3 sentences explaining what the results of these analyses suggest about differences in the FCI (log) between leagues after controlling for the linear and quadratic effects of stadium age, and the effect of cost-of-living? **(2pts.)**

[1] A technical term