Assignment 06

Multilevel Regression Models - Longitudinal Analysis

This assignment is intended to give you experience working with mixed-effects regression models to analyze longitudinal data. Turn in a printed document that includes your responses to each of the questions on this assignment. This assignment is worth 15 points.

Please adhere to the following guidelines for further formatting your assignment:

- All graphics should be resized so that they do not take up more room than necessary and should have an appropriate caption.
- Any typed mathematics (equations, matrices, vectors, etc.) should be appropriately typeset within the document.
- If you are using Markdown, all syntax should be hidden (i.e., not displayed) unless specifically asked for. Any messages or warnings produced from loading packages should also be hidden.

For this assignment, you will use the files, *nhl-fci-by-year.csv* and *nhl-team-metadata.csv*. The first file contains 9 seasons of data on 31 NHL teams. The variables in this file are:

- `team`: NHL team name
- `year`: NHL season (e.g., 2002 indicates the 2002–2003 NHL season)
- `fci`: Fan cost index (FCI) for each season. There are no data for 2012, since that year the NHL was locked out. The FCI comprises the prices of four (4) average-price tickets, two (2) small draft beers, four (4) small soft drinks, four (4) regular-size hot dogs, parking for one (1) car, two (2) game programs and two (2) least-expensive, adult-size adjustable caps. Costs were determined by telephone calls with representatives of the teams, venues and concessionaires. Identical questions were asked in all interviews.

The second file includes team-level metadata for each of the 31 teams. The variables in this file are:

- `team`: NHL team name
- `established`: Year the team was established
- `arena`: Name of the arena where the team plays
- `arena_location`: City and State/Provence where the team's arena is located
- `arena_open`: The year the arena opened
- `arena_capacity`: Seating capacity of the arena
- `avg_temp`: Average temperature for the location of the team's arena
- `latitude`: Latitude of the arena (north/south)
- `longitude`: Longitude of the arena (east/west)
- `canada`: Does the team play in Canada? (0=No, 1=Yes)
- `division`: Team's NHL division
- `conference`: Team's NHL conference

Preparation

After importing both data sets, create a new `time` variable in the fci-by-year data that centers the `year` values at 2002. In other words, the new `time` variable will have a value of 0 when the `year` variable is 2002, a value of 1 when the `year` variable is 2003, etc. This variable will represent the number of years since 2002. Then, merge the `avg_temp` variable from the team metadata into this data.

For all analyses in this assignment, use the `time` variable and not the `year` variable. Also, when fitting mixed-effects models, use maximum likelihood estimation.

Data Exploration

1. Create a spaghetti plot of FCI over time for each team. In this plot, all teams should be in the same panel. Add the mean FCI over time as well. Make the teams' trajectories slightly transparent so that the mean trajectory is easily visible.

2. Compute the mean and standard deviation of FCI conditioned on year (actually use `year` for this question). Present these in a table.

3. What do the plots and numerical summaries suggest about whether there are mean differences in FCI over time? Explain.

Unconditional Model: Fixed-Effects Stucture

Fit two unconditional models : (1) the unconditional means model; and (2) the unconditional linear growth model. In each of these models only include a random-effect for intercept. Label these models, *Model A*, and *Model B* respectively.

4. Include the results from Models A–B in a table of model results. Be sure to include any goodness-of-fit and/or pseudo-R^2 measures for each model that you feel are appropriate. You will be using the results from these three models to determine the structure of the Level-1 model, so be sure to use an appropriate method of estimating the effects. (*Note: You will be adding models to this table later in the assignment, so you may need to add additional rows, etc. into your table.*) **(2pts)**

5. Based on the results of fitting the unconditional models, which fixed-effects structure should be adopted for the Level-1 growth model? Explain.

Unconditional Model: Random-Effects Stucture

6. Create another spaghetti plot of FCI over time facetted by team. In this plot, each team should be in a separate panel. By referring to this plot, indicate whether there is graphical evidence to include an additional random-effects of linear growth (*Note: You do not need to include the plot in the printed document you submit, only your explanation.*)

7. Using the fixed-effects structure you adopted for the Level-1 model in Question #5, fit a model (Model C) that includes random-effects for both intercept and linear growth. Add the results from this model to the table of model results you created in Question #4.

8. Based on the model evidence, adopt one of the following models: Model A, Model B, or Model C. Write the multilevel *and* composite equations for your adopted model. **(2pts)**

Covariate Models

9. Using your adopted structure for the fixed- and random-effects, now also include potential covariates. In Model D, include the fixed-effect of average temperature. Add the results from fitting this covariate model to your table of model results. Also fit a model that includes the fixed-effect of average temperature and the cross-level interaction between average temperatue and linear growth (Model E). Add these results to your table of model results as well. (*Note: This table should now include five models.*) **(2pts)**

10. Compare the unconditional model (Model C) to both potential covariate models (Models D and E). Which should be adopted? Explain.

Plot of Model Results

11. Create a single plot that includes a line displaying the predicted FCI as a function of time for an average team (fixed-effect line). Use the league average temperature to plot this line. Add a line showing the predicted FCI as a function of time for the Minnesota Wild. To plot this line, you will need to use the verage temperature for the Wild and the team-effects for the Wild (fixed- and random-effects line for the Wild). Write a caption for your plot (2–3 sentences) that help a reader understand the effect of time and temperature on FCI. **(2pts)**