

Assignment 06

Mixed-Effects 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. 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 file *nhl.csv* (see the [data codebook](#) to examine longitudinal variation in cost of attending an NHL game. This assignment is worth 15 points.

Preparation

After importing the data set, create a new variable called `time` that centers the `year` values at 2002. In other words, the `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.

For all analyses in this assignment, unless stated otherword, use the `time` variable and not the `year` variable.

Data Exploration

1. Create and report a plot to display the cost of attending an NHL game (FCI) over time for each team (team profiles). In this plot, all teams should be in the same panel. Also add the profile based on the mean cost of attending an NHL game FCI over time. Make the teams' profiles slightly transparent so that the mean profile 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 the cost of attending an NHL game over time? Explain.
4. Does the plot of the teams' profiles, indicate whether models that include a random-effect of linear growth should be included in the set of candidate models? Explain.
5. Create and report a plot of the average profile for the cost of attending an NHL game over time, but this time facet on whether or not high school hockey is a tradition in the team's location.
6. Based on the plot created in Question #5, indicate whether models that include an effect of high school hockey tradition should be included in the candidate set of models. If this effect should be included, should it be a main-effect or an interaction-effect with time? Explain.

Fitting and Evaluating Models

Fit the following five models:

- **Model A:** Unconditional random intercepts model
- **Model B:** Unconditional growth model (with only intercept random-effect)
- **Model C:** Unconditional random growth model (with both intercept and linear growth random-effects)
- **Model D:** Random growth model with main-effect of high school hockey (with both intercept and linear growth random-effects)
- **Model E:** Random growth model with interaction-effect between high school hockey and time (with both intercept and linear growth random-effects)

Use maximum likelihood estimation to fit all the models.

7. Create and report a table of regression results. Include the estimated fixed-effect coefficients and standard errors for each of the five fitted models in this table. Also include the model-level log-likelihood and AICc measures for each of these models. **(2pts)**
8. Create and report a table that includes the estimated variance components for each of the fitted models. **(2pts)**
9. Create and report a table of model evidence that allows you to compare the five fitted models.

Adopted Model

9. Write the multilevel equations for the statistical model (not the fitted equations) for the model with the most empirical evidence.
10. Write the team-specific fitted equation for the Minnesota Wild based on the model that has the most empirical evidence.
11. Use the results from the fitted model with the most empirical evidence to create a plot that displays the predicted average cost of attending an NHL game as a function of time for teams that have a tradition of high school and those that do not.
12. Add a line to the plot you created in Question #12 (don't re-create the plot) showing the predicted cost of attending a Minnesota Wild game. (To add a specific line to your plot see the [documentation for `geom_abline\(\)`](#)).
13. Write a caption for your plot (3–4 sentences) that help a reader understand the effect of time and high school hockey tradition on cost of attending an NHL game and also how the cost of attending a Wild game compares to the average team over time.