Assignment 09

Transformations

Each season, Team Marketing Report (TMR) computes the cost of taking a family of four to a professional sports contest for each of the major sporting leagues. The cost, referred to as *Fan Cost Index* (or FCI) is a summary of what it costs to take a family of four to a game. It comprises the prices of four (4) adult 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. The FCI is determined by telephone calls with representatives of the teams, venues and concessionaires. Identical questions were asked in all interviews. Prices for Canadian teams were converted to US dollars and comparison prices were converted using a recent exchange rate.

In this assignment, you are going to focus on whether or not differences in sporting league predicts variation in the cost of going to a game in the 2014/2015 season. Use the data from the file *fan-cost.csv* to answer the questions on this assignment. Each question is worth one point unless otherwise noted. The entire assignment is worth 14 points.

The file *fan-cost.csv* contains data on the cost of attending a sporting event during the 2014/2015 season for 122 professional sports teams across the United States. There is also data on the const of attending a sporting event for the same teams during the 2013/2014 season. The variables are:

- `team`: Name of professional sports team
- `league`: Major sporting league the team plays in (MLB = Major Lague Baseball; NBA = National Basketball Association; NFL = National Football League; NHL = National Hockey League)
- `fci_2013`: Fan Cost Index (FCI) for 2013/2014 season
- `fci_2014`: Fan Cost Index (FCI) for 2014/2015 season

Please submit your responses to each of the questions below in a printed document. Also, 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 and labels.
- Any typed mathematics (equations, matrices, vectors, etc.) should be appropriately typeset within the document using Equation Editor, Markdown, or LaTeX.

This assignment is worth 14 points. Each question is worth one point unless otherwise noted.

Effect of League on 2014 FCI

1. Create and examine the scatterplot of the relationship between sporting league and 2014 FCI. What does the scatterplot suggest about the tenability of homogeneity of variance? Explain.

To alleviate problems with homogeneity of variance, we will log-transform the 2014 FCI variable using the natural logarithm. (Use this log-transformed outcome in all the remaining analyses.) Also create a dummy variable for each sporting league; four total. Fit the model using league to predict variation in FCI. Make NBA the reference group in this model and in all remaining models you fit in this assignment.

2. Write the equation for the fitted model using Equation Editor (or some other program that correctly typesets mathematical expressions).
3. Report and interpret the NHL coefficient using by back-tranforming FCI.

Effect of 2013 FCI on 2014 FCI

4. Create and examine the scatterplot of the relationship between 2013 FCI and the log-transformed 2014 FCI outcome. Add the loess smoother to this plot. Does this plot suggest that the relationship between 2013 FCI and log-transformed 2014 FCI is linear or non-linear? Explain.
5. Log-transform the 2013 FCI predictor using the natural logarithm. Create and examine the scatterplot of the relationship between log-transformed 2013 FCI predictor and the log-transformed 2014 FCI outcome. Add the loess smoother to this plot. Does this plot suggest that the relationship between log-transformed 2013 FCI and log-transformed 2014 FCI is linear or non-linear? Explain.

Fit the model using log-transformed 2013 FCI to predict variation in log-transformed 2014 FCI.

6. Write the equation for the fitted model using Equation Editor (or some other program that correctly typesets mathematical expressions).
7. Interpret the effect of 2013 FCI on 2014 FCI using the language of percentage change.

Controlled Model to Examine League Differences

Fit the model that includes both league and log-transformed 2013 FCI to predict log-transformed 2014 FCI. Again, use NBA as the reference group.

8. Examine the residuals from the fitted model. Are the assumptions satisfied? Explain.
9. Report and interpret (1) the NHL coefficient using the back-transformed FCI, and (2) the effect of 2013 FCI on 2014 FCI using the language of percentage change.
10. Create a table (suitable for publication) that presents each of the possible pairwise contrasts examining controlled differences between the four leagues. Include the unadjusted p-values, and the Benjamani–Hochberg adjusted p-values for the controlled differences. (Note: To obtain all of these, you may need to fit additional models.) **(2pts.)**
11. Use the table to describe differences in 2014 FCI between leagues after accounting for differences in 2013 FCI.
12. Using the `ggplot()` function, create a plot of the back-transformed fitted model that allows you to visualize the effects of 2013 FCI (back-transformed) and league on 2014 FCI (back-transformed). **(2pts.)**