More than Two Groups

EPsy 8251

Assignment #6

You will use the data set *FCI.csv* to examine whether there are mean differences in the cost of going to a game between the four major professional sports leagues in the United States. In these data, the variable `fci` is the outcome and `league` is the predictor. 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 17 points possible for the assignment (each question is worth one point).

EXAMINE THE SAMPLE EFFECT OF LEAGUE

1. Create a scatterplot to examine whether there are differences in the FCI between the four leagues. Put `league` on the x-axis and `fci` on the y-axis. Make the observations in the plot semi-transparent (so they are not the highlight of the plot). Add points to display the mean FCI for the four leagues. These points should be completely opaque and larger than the points displaying each observation in the data. Finally, connect the mean points with lines to help readers compare the mean values.

2. Compute the mean FCI, the standard deviation, and the sample size for the four leagues. Present these in a table (with an appropriate caption and header row) that allows them to easily be compared.

3. Based on the sample means, what do the data suggest about potential effect of league (i.e., league differences in FCI)? Explain.

EXAMINE THE EFFECT OF LEAGUE: INFERENCE

4. Create appropriate dummy variables so that you can fit a single regression model to test the effect of league. Write the omnibus null hypothesis associated with testing the population "variance accounted for" measure.

5. Fit the regression model using your dummy variables, and use the output to explain what the data suggest about whether differences in league explain variation in FCI for the "population".

6. Using the output from your fitted regression model, write the regression equation for your fitted model.

7. Interpret each of the four regression coefficients from your fitted model. **(2pts.)**

EFFECT SIZE

8. Using the regression output, provide an estimate of η^2.

9. Interpret the computed estimate of η^2.

10. Compute the 95% confidence interval for η^2.

PAIRWISE COMPARISONS

11. Carry out the *post hoc* analysis to examine all of the pairwise league comparisons using the Benjamani–Hochberg adjustments to the p-values. Also obtain the Benjamani–Hochberg adjusted simultaneous intervals. Put these results in a table along with the tested contrast and the estimated contrast value. (Note: This will be similar to the table shown at the end of the *Pairwise Comparison* notes.) Be sure your table has an appropriate caption. **(2pts.)**

12. What do the results of these analyses suggest about differences in the FCI between leagues? Write a few sentences that explain these differences using non-technical language.

13. Create a plot, using `ggplot()`, to display each of the six simultaneous intervals. To see an example of how this plot should look, assign the results from obtaining the simultaneous intervals to an object, and then use the `plot()` function on that object. **(3pts.)**