DATA 5600

Final Project

- Instructor: Tyler J. Brough
- Date Assigned: 04.25.MON
- Due Date: 05.04.WED at Midnight

1. Interpret the OLS regression model in two ways:

- **(a)** As an optimization problem. What is the objective function?
- **(b)** As a classical (frequentist) probability model. Where, when and how does probability apply to the regression problem?

2. Use the HPRICE.XLS data set.

- **(a)** Using OLS estimation perform regression analysis using all the tools that you have learned in this class to comment on the model (i.e. coefficient estimates, R^2, hypothesis testing, confidence intervals, etc).
- **(b)** Redo the above analysis in part (a) using the bootstrap. Compare and constrast with your analysis in part (a). Interpret the bootstrap approach as a probability model. Conduct your bootstrap analysis in two different ways:
  - **i.** A parametric bootstrap
  - **ii.** A nonparametric bootstrap
- **(c)** Redo the regression analysis again using a Bayesian regression model.
  - **i.** Compare and constrast with your results from parts (a) and (b).
  - **ii.** Present $90\%, 95\%$ and $\%99$ highest density intervals (HDIs). Compare and constrast with the confidence intervals that you obtained in parts (a) and (b). Carefully interpret these HDIs. How do they differ from the confidence intervals, etc?
  - **iii.** Conduct prior predictive simulation. Present your results. Comment on the results.
  - **iV.** Conduct posterior predictive simulation. Present your results. Comment on the results.

3. Explain what it means to "think statistically" from the two views of probability that we have discussed this semester. You may want to discuss some of the following (and more!):

- **(a)** Throughout the semester we have drawn a table about which quantities are held fixed and which are considered random from the two different perspectives on probability. Reproduce that table. Compare and constrast. Comment on how this changed your view of what probability is.
- **(b)** How has learning these two views changed your view of data analytics?

- **(c)** Discuss how learning and comparing and constrasting these two views will help you going forward to confront data analytics tasks and to face uncertainty in general.

4. What were some of the most important things you learned this semester? What things will you take with you from this class?