Model selection and confidence intervals

Homework Week 7

The assignment

Assignment

Budget cuts at the Newfoundland-Dachshund Society of Underdogs (NDSU) have forced administration to evaluate how departments are managed, and they have turned to you for answers. Using the Chatterjee–Price Attitude data—available in `R` as `data(attitude)`—as an example, the bosses want you to determine which factor or factors drive employee satisfaction in the surveys, measured as the response variable `rating`. Note that the bosses only want you to focus on those variables determined most relevant to NDSU—complaints, privileges, and learning—since some are right out; for example, raises aren't an option under the new budget climate, and its Minnesota, so no one's going to be too critical der, eh?

Right. So work up the data as one does and apply functions and tests we learned this week. But one rule: **no *P* values!**

Data preparation

Identification

Identify the structure of the data.

Assumptions

Provide graphs on data distribution, fit an appropriate Probability Density Function, and test model assumptions.

Answer these questions:

- What PDF fits the response variable?
- What statistical model will you apply? What are its assumptions, and how do you know the data fit those assumptions?

Model fitting and selection

Define model set

1. Construct a set of candidate models
2. Store the models and their names in a candidate model list
3. Report a table comparing the models, sorted by an appropriate information criterion.

Model selection

Answer these questions about the model set:

- Which model or models do you consider competitive? Based on what?

- What do these models have in common?

Model averaging

Calculate and report a single parameter estimate, and associated 95% confidence interval, for each term included in the list of models you determined above to be competitive.

Plot confidence intervals

Plot parameter estimate(s) and associated 95% confidence interval(s) in a single graph.

Text answer:

What does it mean to say confidence intervals report an *effect size*?

Conclusions

Give your recommendation on which variable(s) most influence employee satisfaction ratings. In your response include magnitude of effect, sign (whether the effect is positive or negative), and compare the effect to other variables, if present.