

# Nonparametric Regression Exercise

February 28, 2020

1. Using the Stata dataset <http://www.stata-press.com/data/r13/motorcycle>, explore the relationship between time and acceleration
  - Start with the Stata default settings for `lpoly`
  - What is the impact of picking:
    - Different polynomial degrees?
    - Different bandwidths?
    - Different kernel functions?

Produce one graph for each of i) to iii) with three choices for each criteria to show what happens as they are varied.
2. Estimate the regression with a bandwidth of 10. A researcher complains that the “bandwidth is too big”. What does this mean? What are the implications for the regression estimate? How would you optimally select the bandwidth?
3. What problems might you encounter when predicting out-of-sample with nonparametric methods?
4. Why does the confidence interval vary with different values of  $x$ ?