Applied Regression

Dr. Frank Wood

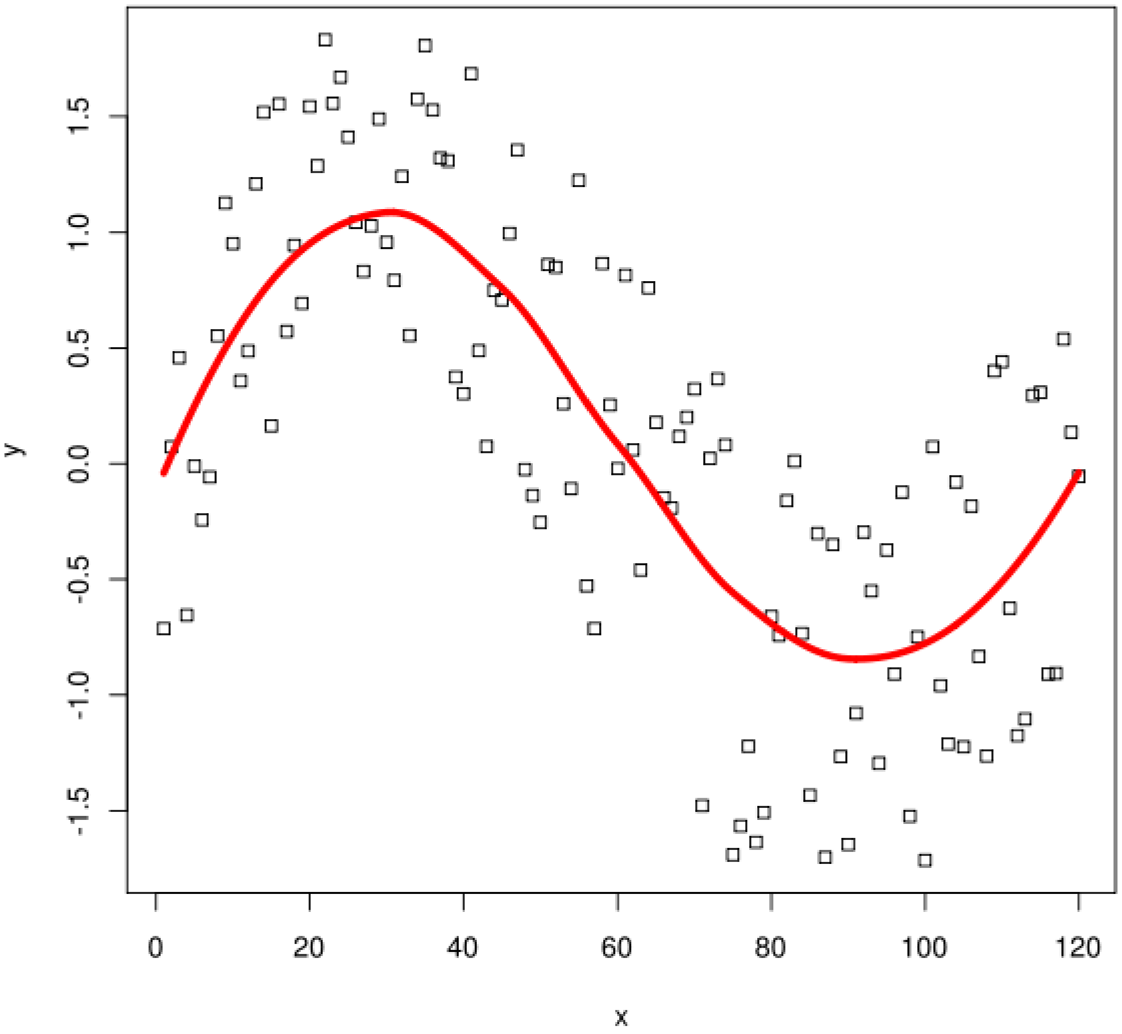
Nonparametric Regression Curves

- So far we have only dealt with parametric regression approaches
  - Linear
  - Transformed
- Other approaches exist as well
  - Method of moving averages
    - Interpolate between means outputs at adjacent inputs
  - Lowess
    - Locally weighted scatterplot smoothing

Lowess Method

- Intuition
  - Fit low-order polynomial (linear) regression models to points in a neighborhood
    - The neighborhood size is a parameter
    - Determining the neighborhood is done via a nearest neighbors algorithm
  - Produce predictions by weighting the regressors by how far the set of points used to produce the regressor is from the input point for which a prediction is wanted
- While somewhat ad-hoc, it is a method of producing a nonlinear regression function for data that might seem otherwise difficult to regress.

Lowess example



Bonferroni Joint Confidence Intervals

- Calculation of Bonferroni joint confidence intervals is a general technique
- In class we will highlight its application in the regression setting
  - Joint confidence intervals for β_0 and β_1
- Intuition
  - Set each statement confidence level to greater than 1-α so that the family coefficient is at least 1-α

Ordinary Confidence Intervals

- Start with ordinary confidence limits for β_0 and β_1

$$\begin{aligned} b_0 &\pm t(1-\alpha/2; n-2)S\{b_0\} \\ b_1 &\pm t(1-\alpha/2; n-2)S\{b_1\} \end{aligned}$$

- And ask what the probability that one or both of these intervals is incorrect.

General Procedure

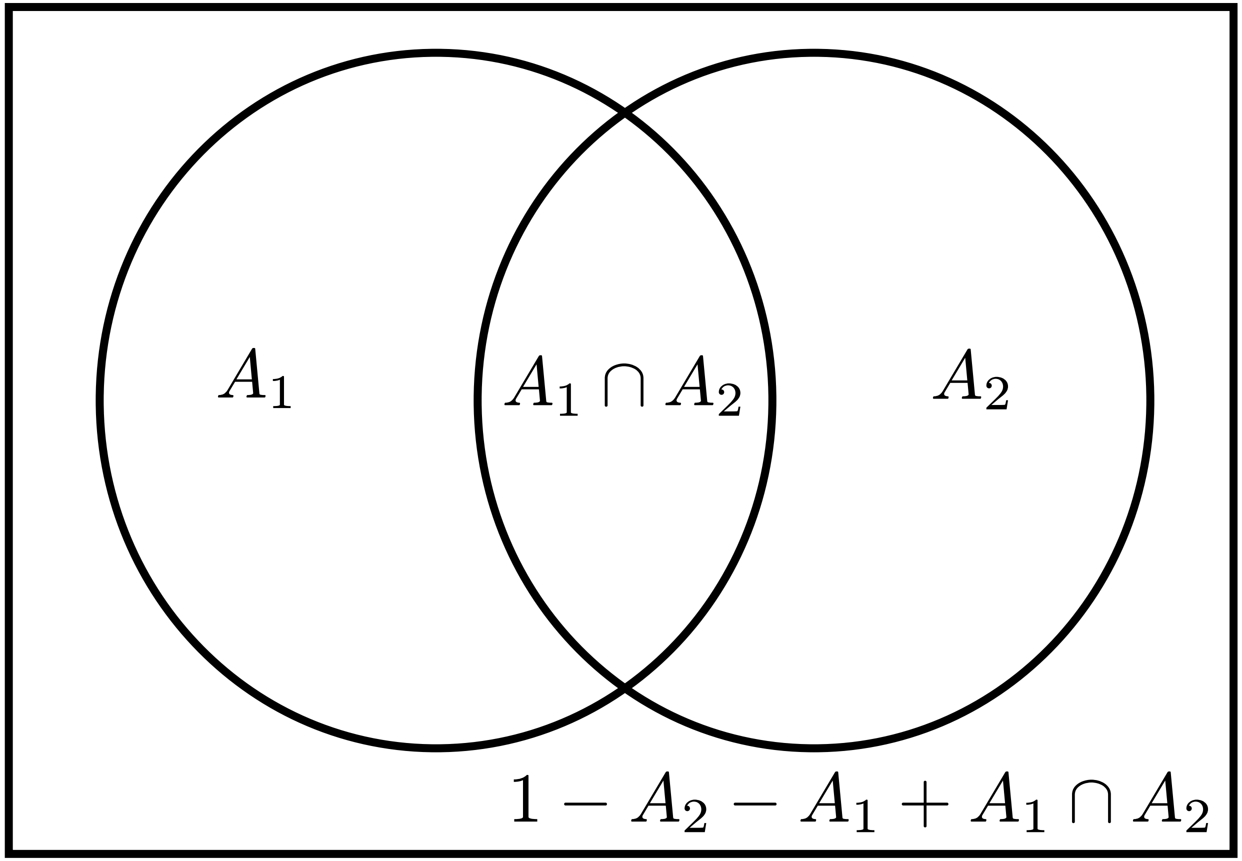
- Let A_1 denote the event that the first confidence interval does not cover β_0
- Let A_2 denote the event that the second confidence interval does not cover β_1.
- We want to know

$$P(\bar{A}_1 \cap \bar{A}_2)$$

- We know

$$P(A_1) = \alpha \;\; P(A_2) = \alpha$$

Venn Diagram



- Bonferroni inequality

$$P(\bar{A}_1 \cap \bar{A}_2) \geq 1 - P(A_1) - P(A_2)$$

Joint Confidence Intervals

- If we know that β_0 and β_1 are estimated with, for instance, 95% confidence intervals, the Bonferroni inequality guarantees us a family confidence coefficient of at least 90% (if both intervals are correct)

$$P(\bar{A}_1 \cap \bar{A}_2) \geq 1 - \alpha - \alpha = 1 - 2\alpha$$

- To pick a family confidence interval (bound) the Bonferroni procedure instructs us how to adjust the value of α for each interval to achieve the overall interval of interest

1-α family confidence intervals

- … for β_0 and β_1 by the Bonferroni procedure are

$$b_0 \pm Bs\{b_0\} \; b_1 \pm Bs\{b_1\}$$

$$B = t(\boxed{1 - \alpha/4}; n - 2)$$

Misc. Topics

- Simultaneous Prediction Intervals for New Observations
  - Bonferroni again
- Regression Through Origin
  - One fewer parameters
- Measurement errors in X
- Inverse Prediction