Nonparametric Regression

Johns Hopkins Engineering

625.461 Statistical Models and Regression

Module 14 – Lecture 14D

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

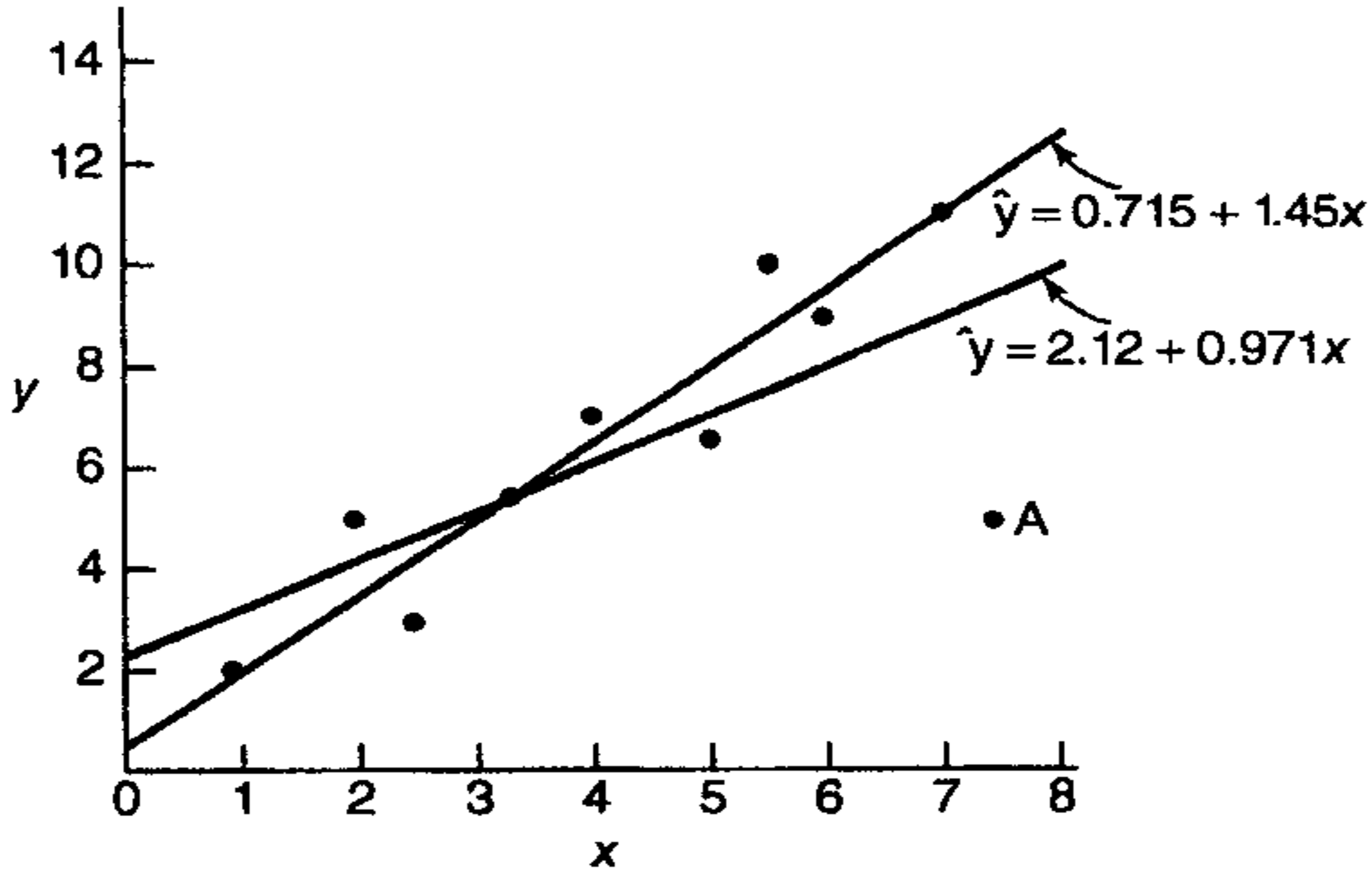
Nonparametric Regression and Robust Estimation

y: response variable

x: regressor variables

$$y_i = x_i{}'\beta + \varepsilon_i \qquad i = 1, \ldots, n$$

Impact of an Influential Observation



Nonparametric Regression and Robust Estimation

One better way of regression analysis is to define a class of robust (or M-) estimators:

$$\underset{\boldsymbol{\beta}}{\text{Minimize}} \sum_{i=1}^{n} \rho(e_i) = \underset{\boldsymbol{\beta}}{\text{Minimize}} \sum_{i=1}^{n} \rho(y_i - \mathbf{x}_i'\boldsymbol{\beta})$$

or a scale-invariant version

$$\underset{\boldsymbol{\beta}}{\text{Minimize}} \sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = \underset{\boldsymbol{\beta}}{\text{Minimize}} \sum_{i=1}^{n} \rho\left(\frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}}{s}\right)$$

where s is a robust estimate of scale.

Nonparametric Regression and Robust Estimation

Operationally, set the first-order partial derivative of ρ with respect to β's to zero and solve for β's.

$$\sum_{i=1}^{n} x_{ij}\,\psi\left(\frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}}{s}\right) = 0, \quad j = 0, 1, \ldots, k$$

Nonparametric Regression and Robust Estimation

In some other scenarios, performing regression on ranked data or rank-scored data may be a viable alternative to performing regression on original data.

That is, rank y or rank x or both before regression

This can dampen the influence of outliers or extreme values.

JOHNS HOPKINS

WHITING SCHOOL *of* ENGINEERING