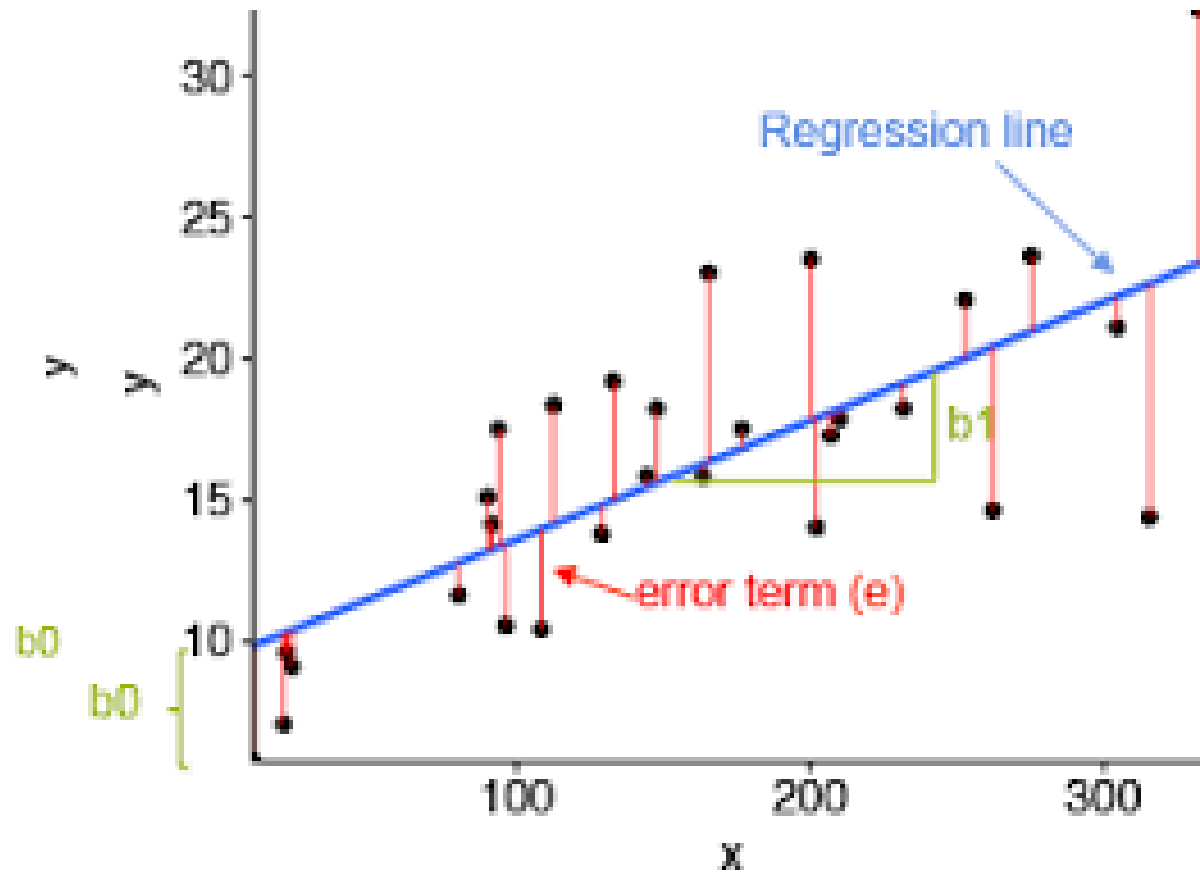


Sports Econometrics

Regression Formation and Reading Results

The only Theory you'll need

- $Y = XB + e$
 - Y is the dependent variable (dependent on $XB + e$)
 - This is the column of data you want to predict
 - X are the independent variable(s)
 - This is the column(s) of data that are predicting Y
 - B are the coefficients on the independent variables
 - B tells you how much X matters
 - e is the error term



A regression minimizes the SUM of squared errors
 $(\text{SUM } e^2 = (Y - XB)^2)$

Let's do a regression

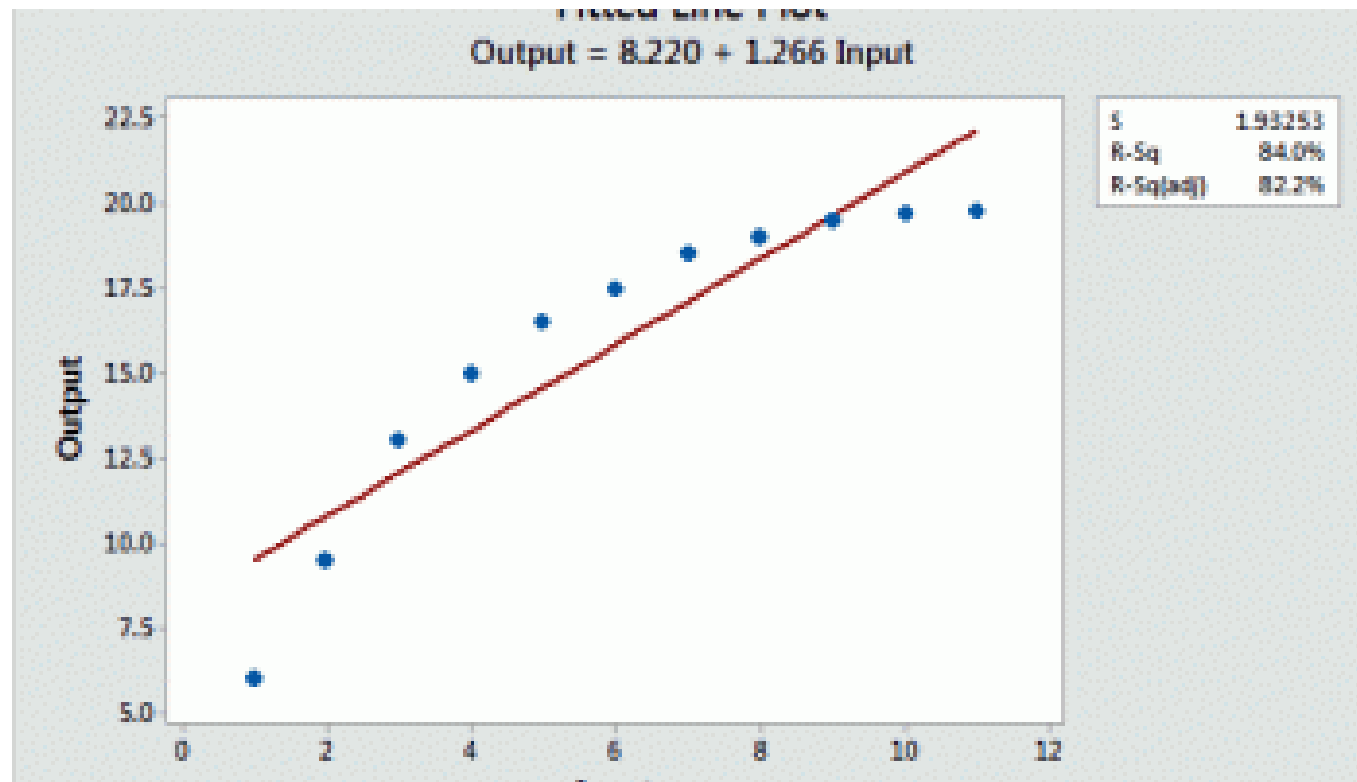
- Excel o'clock

What just happened?

- R-squared
 - How well does XB explain Y? 0%-100%
- Coefficient
 - Slope of X impacting Y
- Standard Error
 - estimate of the *standard deviation* of the coefficient
- T-stat (1.645=10%, 1.960=5%, 2.326=1%)

New Variables, Dummy Variables

- Creating new variables in Excel
- Linear and non-linear effects



Non-linear fitting in Regression

- Easy to interpret
 - Change variable x into $\text{LN}(X)$
- Less easy to interpret
 - Include x^2 and x

Dummy (indicator) variables

- Dummy variables are 0 or 1
 - If condition is true then Dummy=1, Else=0
- Excel and the IF Statement
 - Sets and IF/THEN Condition
 - IF(Condition,ValueIFTRUE,ValueIFFALSE)
 - Stacking your IFs

Stacking your IFs

- Suppose you want data that satisfies: TWO Conditions.
- $\text{IF}(\text{ConditionA}=\text{True}, \text{IF}(\text{ConditionB}=\text{TRUE}, 1, 0), 0)$
- This dummy variable is 1 only if both conditions are met. Otherwise=0

Small Samples

- The Degrees of Freedom problem
 - YOU MUST HAVE 30 DEGREES
 - Why? Law of Large Numbers needs to work
- How to calculate Degrees of Freedom?
 - $\text{DoF} = \# \text{ of Observations} - \# \text{ of X variables}$
 - With small samples you must constrain your # of X variables
 - How? Be picky OR take small steps

Too many X's

- Categorical Variables can't all be included
- Need a null variable
- Categorical variables are relative to null variable
 - Conference Dummy Example
 - Excel informs you that you're wrong

Excel time!

- Let's experiment with non-linear fits
- Let's experiment with making/testing dummies
- Let's stack some IFs and experiment