10/6 Linear Regression

Tuesday, October 6, 2020 11:14 AM

Simple Linear Regression

Determining the relationship between two variables. "Explain Y in terms of X"

Example:

How Sales Vary w/ Advertising
Variation of Crop Yield w/ Amount of Fertilizer
Salary vs Number of Years of Education

Issues:

- Other factors need to be considered
- Define functional form
- How to capture ceteris paribus (all other factors remain the same) relation b/w x and y

Alternative Titles for…

X Variables	Y Variables
Independent Variable	Dependent Variable
Explanatory	Explained
Control	Response
Predictor	Predicted
Regressor	Regressand

Functional Form of Linear Regression

$y = \beta_0 + \beta_1 x + \varepsilon$
$x =$ Independent Variable
$y =$ Dependent Variable

Example

- $Sales = \beta_0 + \beta_1(Advertising) + \varepsilon$
  If we are putting a certain amount of money on advertising, on average, how much do we get back in sales
  $\varepsilon =$ All other factors that impact sales and are not accounted for in X
- $Crop\ Yield = \beta_0 + \beta_1(Fertilizer) + \varepsilon$
  ε captures land quality, rainfall, etc…
- $Salary = \beta_0 + \beta_1(Years\ of\ Education) + \varepsilon$
  ε captures innate ability, experience, work ethic, #of years w/ current employer, etc…

β_1 - Slope
β_1 measures change in y for one unit change in x when all other factors are held constant
β_0 - Intercept
β_0 is the value of why when X = 0

Assumptions

1. Y | X is normally distributed
2. Average value of ε in the population is zero
   Therefore, unobserved factors affecting the response variable has zero average in the entire population
3. ε and x are not correlated

10/6 Linear Regress Page 2

Tuesday, October 6, 2020 12:09 PM

Correlation

Varies between -1 and +1

-1 = perfectly -vely correlated

…

$$\gamma_{xy} = \frac{\varepsilon(x-\bar{x})(y-\bar{y})}{\sqrt{\varepsilon(x-\bar{x})^2}\sqrt{\varepsilon(y-\bar{y})^2}}$$

Zero Conditional Mean Assumption

$E(\varepsilon|x) = E(\varepsilon) = 0$

This assumption defines your β_0

- $Salary = \beta_0 + \beta_1(Years\ of\ Education) + \varepsilon$

ε = inherent ability to work

Therefore, ability is the same, regardless of years of education

Therefore, $E(ability \mid 5\ years\ of\ education) = E(Ability \mid 15\ years\ of\ education)$

Therefore, average ability is the same for all education levels

- $E(Y \mid X) = \beta_0 + \beta_1 X + \varepsilon$

On average, how much does Y change w/ X

Ex: Y = 1.5 + 0.5X

Y = College GPA

X = High School GPA

  - For a student who has a 4.0 GPA in high school has a college of:

    Y = 1.5 + 0.5(4.0)

    Y = 3.5