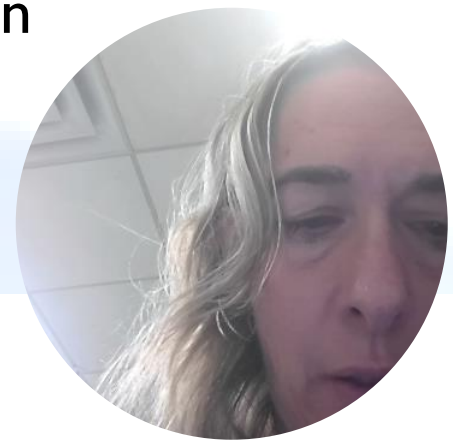
V39 - Multiple Linear Regression - Part 1

- Introduction to Multiple Linear Regression
- Multicollinearity

Course: Statistical Testing & Regression
Dr. Renee Clark
Swanson School of Engineering
Industrial Engineering
University of Pittsburgh



Regression Analysis

- Simple Linear Regression
  - 1 independent variable
  - Need a foundation!
  - Reasonable at times
- Many scientific problems or real-world applications tend to be more complex
  - Require more than 1 independent variable to explain variation & make good predictions
    - multiple **regression model**
    - > 1 independent variable

Multiple Linear Regression Model

$$y = \beta_0 + \beta_1 x + \varepsilon$$

simple linear regression
SLR

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_k x_k + \varepsilon$$

MLR

"k" indep variables

$$\varepsilon \sim N(0, \sigma^2)$$

- Each β_i estimated via Least Squares Method → obtain b_i's

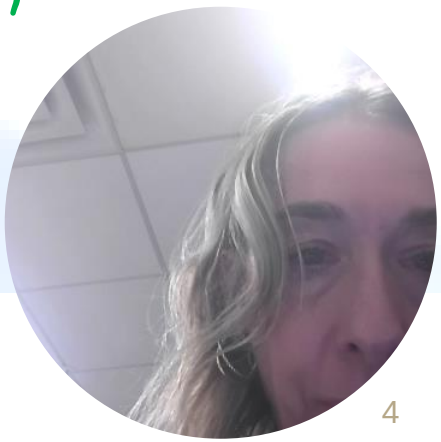
$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \ldots + b_k x_k$$

Multicollinearity

- **Multicollinearity** refers to presence of linear relationships between *independent variables*

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_k x_k + \varepsilon$$

- Investigate **multicollinearity** via correlation coefficient (r)



Multicollinearity

Correlation matrix for *4 independent variables: x1, x2, x3, x4*

Correlation: x1, x2, x3, x4

	x1	x2	x3
x2	0.952		
	0.000		
x3	0.534	0.263	
	0.139	0.495	
x4	0.390	0.155	0.784
	0.299	0.691	0.012

Cell Contents: Pearson correlation
P-Value

Substantial correlations ($p < 0.05$):

X1 & X2
X3 & X4

$H_0: \rho = \phi$ X
$H_1: \rho \neq \phi$ ✓

Acknowledgement

This material is based upon work partially supported by the National Science Foundation under Grant# 2335802. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

You are free

- to **Share** – to copy, distribute, display and perform the work
- to **Remix** – to make derivative works

Under the following conditions

- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Noncommercial** — You may not use this work for commercial purposes.
- **Share Alike** — If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

THE END