

# Regression Analysis

## Multiple Linear Regression

**Nicoleta Serban, Ph.D.**

*Professor*

School of Industrial and Systems Engineering

Model Interpretation



1

## About This Lesson



2

1

# Model Interpretation: Parameters

The Least Squares estimated coefficients have specific interpretations:

- $\hat{\beta}_0$  The estimated expected value of the response variable when all predicting variables equal zero.
- $\hat{\beta}_i$  The estimated expected change in the value of the response variable associated with one unit of change in the value of the  $i^{\text{th}}$  predicting variable (i.e., associated with a one-unit change in  $x_i$ , where  $i$  is any of  $1, \dots, p$ ), holding all other predictors in the model fixed (i.e., holding fixed  $x_j$  for  $j = 1, \dots, p$  where  $j \neq i$ ).



3

# Model Interpretation: Simple vs. Multiple Regression

**Marginal** versus **conditional** relationship:

- Marginal** Simple linear regression captures the association of a predicting variable to the response variable marginally, i.e., *without consideration of other factors*.
- Conditional** Multiple linear regression captures the association of a predicting variable to the response variable conditionally, i.e., *conditional of all other predicting variables in the model*.

The estimated regression coefficients for conditional and marginal relationships can differ not only in magnitude but also in sign or direction of the relationship.



4

# Model Interpretation: Causality vs. Association

**Causality** Statements: Experimental Designs

**Association** Statements: Observational Studies

**Example:** We take a sample of college students and determine their college grade point average (*COLGPA*), high school GPA (*HSGPA*), and SAT score (*SAT*). The estimated model is:  $COLGPA = 1.3 + 0.7(HSGPA) - 0.0003(SAT)$ .

- **Incorrect Interpretation:** Higher values of SAT are associated with lower values of College GPA.
- **Correct Interpretation:** Higher values of SAT are associated with lower values of college GPA, *holding high school GPA fixed*.

The coefficients of a multiple regression must not be interpreted marginally!



5

## Different Roles of Predicting Variables

Predicting variables can be distinguished as:

- **Controlling** – to control for bias selection in the sample. They are used as ‘default’ variables in order to capture more meaningful relationships.
- **Explanatory** – to explain variability in the response variable. They may be included in the model even if other “similar” variables are in the model.
- **Predictive** – to best predict variability in the response regardless of their explanatory power.



6

# About This Lesson

