Econ 103: Multiple Linear Regression I

Manu Navjeevan

UCLA

July 28, 2021

Content Outline

The Model:

- Adding more covariates
- Assumptions needed for inference

The Estimator:

- Relation to Single Linear Regression Estimator
- Asymptotic Dsitribution

Inference:

- Hypothesis Tests and Linear Combinations
- Confidence Inervals

Modeling Choices:

- Polynomial Equations, transformations, and interactions
- R^2 and goodness of fit

Table of Contents

The Model

The Estimator

Inference

Modeling Choices

The Model: Introduction

So far we have used the model $Y = \beta_0 + \beta_1 X + \epsilon$ defined by the line of best fit parameters

$$\beta_0, \beta_1 = \arg\min_{\tilde{\beta}_0, \tilde{\beta}_1} \mathbb{E}\left[\left(Y - \tilde{\beta}_0 - \tilde{\beta}_1 X\right)^2\right].$$

to learn about the relationship between a single random variable X and Y and to use X to predict Y.

So far we have used the model $Y = \beta_0 + \beta_1 X + \epsilon$ defined by the line of best fit parameters

$$\beta_0, \beta_1 = \arg\min_{\tilde{\beta}_0, \tilde{\beta}_1} \mathbb{E}\left[\left(Y - \tilde{\beta}_0 - \tilde{\beta}_1 X\right)^2\right].$$

to learn about the relationship between a single random variable X and Y and to use X to predict Y.

Examples:

- Using education to predict income or interpreting the coeffecient $\hat{\beta}_1$ to learn about the relationship between the two.

The Model: Introduction

So far we have used the model $Y = \beta_0 + \beta_1 X + \epsilon$ defined by the line of best fit parameters

$$\beta_0, \beta_1 = \arg\min_{\tilde{\beta}_0, \tilde{\beta}_1} \mathbb{E}\left[\left(Y - \tilde{\beta}_0 - \tilde{\beta}_1 X\right)^2\right].$$

to learn about the relationship between a single random variable X and Y and to use X to predict Y.

Examples:

- Using education to predict income or interpreting the coeffecient $\hat{\beta}_1$ to learn about the relationship between the two.
- Learning about the relationship between smoking and heart disease.

The Model: Introduction

However, what happens if we have access to multiple explanatory variables $X_1, \ldots, X_p$?

Examples:

- Suppose we wanted to impact the joint effect of education <u>and</u> experience on age?
- Learn about the relationship between smoking, genetic risk, and heart disease

The Model: Introduction

As before, we may be interested in the parameters of a "line of best fit" between Y and our explantory variables $X_1, \ldots, X_p$:

$$\beta_0, \beta_1, \ldots, \beta_p = \arg \min_{b_0, \ldots, b_p} \mathbb{E}\left[(Y - b_0 - b_1 X_1 - b_2 X_2 - \cdots - b_p X_p)^2\right].$$

The Model: Introduction

As before, we may be interested in the parameters of a “line of best fit” between Y and our explantory variables $X_1, \ldots, X_p$:

$$\beta_0, \beta_1, \ldots, \beta_p = \arg \min_{b_0, \ldots, b_p} \mathbb{E}\left[(Y - b_0 - b_1 X_1 - b_2 X_2 - \cdots - b_p X_p)^2\right].$$

Again defining $\epsilon = Y - \beta_0 - \beta_1 X_1 - \cdots - \beta_p X_p$ these parameters generate the linear model

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \epsilon$$

where, by the first order conditions for β, $\mathbb{E}[\epsilon] = \mathbb{E}[\epsilon X_j] = 0$ for all $j = 0, 1, \ldots, p$.

The Model: Introduction

Example 1: Let Y be log wages, EDU be years of college education, and EXP be years of experience. Prior to this we have estimated the equation

$$Y = \beta_0 + \beta_1 EDU + \epsilon. \tag{1}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 EDU + \beta_2 EXP + \epsilon. \tag{2}$$

The Model: Introduction

Example 1: Let Y be log wages, EDU be years of college education, and EXP be years of experience. Prior to this we have estimated the equation

$$Y = \beta_0 + \beta_1 EDU + \epsilon. \tag{1}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 EDU + \beta_2 EXP + \epsilon. \tag{2}$$

Note that β_0, β_1 in model (1) will differ from β_0, β_1 in model (2).

The Model: Introduction

Example 1: Let Y be log wages, EDU be years of college education, and EXP be years of experience. Prior to this we have estimated the equation

$$Y = \beta_0 + \beta_1 EDU + \epsilon. \tag{1}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 EDU + \beta_2 EXP + \epsilon. \tag{2}$$

Note that β_0, β_1 in model (1) will differ from β_0, β_1 in model (2).

- In (1) β_0 corresponds to the average log wage for someone with no college education

The Model: Introduction

Example 1: Let Y be log wages, EDU be years of college education, and EXP be years of experience. Prior to this we have estimated the equation

$$Y = \beta_0 + \beta_1 EDU + \epsilon. \tag{1}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 EDU + \beta_2 EXP + \epsilon. \tag{2}$$

Note that β_0, β_1 in model (1) will differ from β_0, β_1 in model (2).

- In (1) β_0 corresponds to the average log wage for someone with no college education
- In (2) β_0 will correspond to the average log wage for someone with no college education and no experience

The Model: Introduction

Example 1: Let Y be log wages, EDU be years of college education, and EXP be years of experience. Prior to this we have estimated the equation

$$Y = \beta_0 + \beta_1 EDU + \epsilon. \tag{1}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 EDU + \beta_2 EXP + \epsilon. \tag{2}$$

Note that β_0, β_1 in model (1) will differ from β_0, β_1 in model (2).

- In (1) β_0 corresponds to the average log wage for someone with no college education
- In (2) β_0 will correspond to the average log wage for someone with no college education and no experience
- In (1) β_1 corresponds to the expected change in log wage for an additional year of college education

The Model: Introduction

Example 1: Let Y be log wages, EDU be years of college education, and EXP be years of experience. Prior to this we have estimated the equation

$$Y = \beta_0 + \beta_1 EDU + \epsilon. \tag{1}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 EDU + \beta_2 EXP + \epsilon. \tag{2}$$

Note that β_0, β_1 in model (1) will differ from β_0, β_1 in model (2).

- In (1) β_0 corresponds to the average log wage for someone with no college education
- In (2) β_0 will correspond to the average log wage for someone with no college education and no experience
- In (1) β_1 corresponds to the expected change in log wage for an additional year of college education
- In (2) β_1 corresponds to the expected change in log wage for an additional year of college education <u>after</u> controlling for years of experience

The Model: Introduction

Example 2: Let Y be the (log) final sales price of a home, $SQFT$ be the square footage of the house, and $DAYS$ be the number of days the house has been on the market. Before we estimated and interpreted the linear model:

$$Y = \beta_0 + \beta_1 SQFT + \epsilon. \tag{3}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 SQFT + \beta_2 DAYS + \epsilon \tag{4}$$

The Model: Introduction

Example 2: Let Y be the (log) final sales price of a home, $SQFT$ be the square footage of the house, and $DAYS$ be the number of days the house has been on the market. Before we estimated and interpreted the linear model:

$$Y = \beta_0 + \beta_1 SQFT + \epsilon. \tag{3}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 SQFT + \beta_2 DAYS + \epsilon \tag{4}$$

- In (3) β_0 is interpreted as the average log sales price for a home with zero square feet (an empty lot) regardless of how long it's been on the market.

The Model: Introduction

Example 2: Let Y be the (log) final sales price of a home, $SQFT$ be the square footage of the house, and $DAYS$ be the number of days the house has been on the market. Before we estimated and interpreted the linear model:

$$Y = \beta_0 + \beta_1 SQFT + \epsilon. \tag{3}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 SQFT + \beta_2 DAYS + \epsilon \tag{4}$$

- In (3) β_0 is interpreted as the average log sales price for a home with zero square feet (an empty lot) <u>regardless</u> of how long it's been on the market.
- In (4) β_0 is interpreted as the average log sales price for a home with zero square feet that has just entered the market

The Model: Introduction

Example 2: Let Y be the (log) final sales price of a home, $SQFT$ be the square footage of the house, and $DAYS$ be the number of days the house has been on the market. Before we estimated and interpreted the linear model:

$$Y = \beta_0 + \beta_1 SQFT + \epsilon. \tag{3}$$

Now, we will consider estimation and inference on the model

$$Y = \beta_0 + \beta_1 SQFT + \beta_2 DAYS + \epsilon \tag{4}$$

- In (3) β_0 is interpreted as the average log sales price for a home with zero square feet (an empty lot) regardless of how long it's been on the market.
- In (4) β_0 is interpreted as the average log sales price for a home with zero square feet that has just entered the market
- In (4) β_1 is interpreted as the average change in sales price for a one unit increase in square footage, holding the number of days on the market constant

The Model: Introduction

Example 3: Finally, let's return to an example from Week 1. Let Y be a measure of anxiety levels, ENG be the number of energy drinks consumed per day, and CLS be the number of courses being taken. Before we may have estimated the model:

$$Y = \beta_0 + \beta_1 ENG + \epsilon \tag{5}$$

Now, we may consider the model

$$Y = \beta_0 + \beta_1 ENG + \beta_2 CLS + \epsilon \tag{6}$$

The Model: Introduction

Example 3: Finally, let's return to an example from Week 1. Let Y be a measure of anxiety levels, ENG be the number of energy drinks consumed per day, and CLS be the number of courses being taken. Before we may have estimated the model:

$$Y = \beta_0 + \beta_1 ENG + \epsilon \tag{5}$$

Now, we may consider the model

$$Y = \beta_0 + \beta_1 ENG + \beta_2 CLS + \epsilon \tag{6}$$

- In (5) we can interpret β_0 as the average anxiety level for someone who drinks no energy drinks

The Model: Introduction

Example 3: Finally, let's return to an example from Week 1. Let Y be a measure of anxiety levels, ENG be the number of energy drinks consumed per day, and CLS be the number of courses being taken. Before we may have estimated the model:

$$Y = \beta_0 + \beta_1 ENG + \epsilon \tag{5}$$

Now, we may consider the model

$$Y = \beta_0 + \beta_1 ENG + \beta_2 CLS + \epsilon \tag{6}$$

- In (5) we can interpret β_0 as the average anxiety level for someone who drinks no energy drinks
- In (6) we can interpret β_0 as the average anxiety level for someone who drinks no energy drinks and takes no classes

The Model: Introduction

Example 3: Finally, let's return to an example from Week 1. Let Y be a measure of anxiety levels, ENG be the number of energy drinks consumed per day, and CLS be the number of courses being taken. Before we may have estimated the model:

$$Y = \beta_0 + \beta_1 ENG + \epsilon \tag{5}$$

Now, we may consider the model

$$Y = \beta_0 + \beta_1 ENG + \beta_2 CLS + \epsilon \tag{6}$$

- In (5) we can interpret β_1 as the expected change in anxiety levels for someone who drinks one more energy drink per day

The Model: Introduction

Example 3: Finally, let's return to an example from Week 1. Let Y be a measure of anxiety levels, ENG be the number of energy drinks consumed per day, and CLS be the number of courses being taken. Before we may have estimated the model:

$$Y = \beta_0 + \beta_1 ENG + \epsilon \tag{5}$$

Now, we may consider the model

$$Y = \beta_0 + \beta_1 ENG + \beta_2 CLS + \epsilon \tag{6}$$

- In (5) we can interpret β_1 as the expected change in anxiety levels for someone who drinks one more energy drink per day
- In (6) we can interpret β_1 as the expected change in anxiety levels for an additional energy drink <u>holding the number of courses being taken constant.</u>

The Model: Introduction

Example 3: Finally, let's return to an example from Week 1. Let Y be a measure of anxiety levels, ENG be the number of energy drinks consumed per day, and CLS be the number of courses being taken. Before we may have estimated the model:

$$Y = \beta_0 + \beta_1 ENG + \epsilon \tag{5}$$

Now, we may consider the model

$$Y = \beta_0 + \beta_1 ENG + \beta_2 CLS + \epsilon \tag{6}$$

- In (5) we can interpret β_1 as the expected change in anxiety levels for someone who drinks one more energy drink per day
- In (6) we can interpret β_1 as the expected change in anxiety levels for an additional energy drink <u>holding the number of courses being taken constant.</u>

Question: How may we expect the signs/magnitutes of the parameters to change when going from model (5) to model (6)?

The Model: Questions

Questions?

Estimation: Introduction

Before, in single linear regression when we were interested in the population line of best fit parameters

$$\beta_0, \beta_1 = \arg\min_{b_0, b_1} \mathbb{E}\left[(Y - b_0 - b_1 X)^2\right],$$

we estimated them by finding the line of best fit through our sample $\{Y_i, X_i\}_{i=1}^n$:

$$\hat{\beta}_0, \hat{\beta}_1 = \arg\min_{b_0, b_1} \frac{1}{n} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2.$$

Estimation: Introduction

Before, in single linear regression when we were interested in the population line of best fit parameters

$$\beta_0, \beta_1 = \arg\min_{b_0, b_1} \mathbb{E}\left[(Y - b_0 - b_1 X)^2\right],$$

we estimated them by finding the line of best fit through our sample $\{Y_i, X_i\}_{i=1}^n$:

$$\hat{\beta}_0, \hat{\beta}_1 = \arg\min_{b_0, b_1} \frac{1}{n} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2.$$

$\rightarrow$ Have to estimate these parameters using the sample because we don't know the population distribution of (Y, X)

Estimation: Introduction

Now, we are interested in

Table of Contents

The Model

The Estimator

Inference

Modeling Choices

Table of Contents

The Model

The Estimator

Inference

Modeling Choices

Table of Contents

The Model

The Estimator

Inference

Modeling Choices