

# Logistic Regression

BUS 735: Business Decision Making and Research

## 1

### Goals of this section

- Specific goals:
  - Learn how to conduct regression analysis with a dummy independent variable.
- Learning objectives:
  - LO2: Be able to construct and use multiple regression models (including some limited dependent variable models) to construct and test hypotheses considering complex relationships among multiple variables.
  - LO6: Be able to use standard computer packages such as R to conduct statistical analysis.
  - LO7: Have a sound familiarity of various statistical and quantitative methods in order to be able to approach a business decision problem and be able to select appropriate methods to answer the question.

### Logistic Regression

**Logistic Regression:** method for estimating a regression with a *dummy dependent variable*.

### Examples

- Will a potential customer purchase a product?(YES=1, NO=0).
  - Might use explanatory variables: age, gender, income, etc.
- Will a potential employee be retained after one year?(YES=1, NO=0).
  - Might use explanatory variables: age, gender, years experience, past income, education dummy (4-year = 1, otherwise = 0).

### Regular Regression Model?

- Dependent variable is not interval: variance of residual depends on  $y = 0$  or  $1$ .
- Linear Probability Model: A normal regression with variance correction

### Structural Form

#### Normal regression

$$y_i = b_0 + b_1X_{1,i} + b_2X_{2,i} + \dots + b_{k-1}X_{k-1,i} + e_i$$

#### Logistic regression

$$\log(Odds) = b_0 + b_1X_{1,i} + b_2X_{2,i} + \dots + b_{k-1}X_{k-1,i} + e_i$$

$$Odds = \frac{P(y_i = 1)}{1 - P(y_i = 1)} = \frac{P(y_i = 1)}{P(y_i = 0)}$$

### Predicted Value

#### Normal regression

$$\hat{y}_i = b_0 + b_1X_{1,i} + b_2X_{2,i} + \dots + b_{k-1}X_{k-1,i}$$

#### Logistic regression

Predicted *Logist* or *Log Odds Ratio*:

$$\hat{l}_i = \ln(\widehat{Odds}) = b_0 + b_1X_{1,i} + b_2X_{2,i} + \dots + b_{k-1}X_{k-1,i}$$

### Predicted probability:

$$P(\widehat{y}_i = 1) = \frac{1}{1 + e^{(-\hat{l}_i)}}$$

### Marginal Effects

- **Marginal effect for regression:** measure of how much  $y$  changes when  $x$  increases by 1.
  - Example: How much does public expenditure per capita increase (or decrease) when economic ability increases by one unit?
- **Marginal effect for logit:** measure of how much  $P(y_i = 1)$  changes when  $x$  increases by 1.
  - How much more (or less) likely will an interview candidate be working here after one year if she/he has a four year college education?

## Coefficients versus Marginal Effects

### Regular regression: Coefficient $b_2$

- The sign (positive/negative) indicates whether  $x_2$  causes  $y$  to increase or decrease.
- The magnitude tells *how much*  $y$  increases when increasing  $x_2$  by 1.
- Coefficient = Marginal Effect.

### Logistic regression: Coefficient $b_2$

- The sign (positive/negative) indicates whether  $x_2$  causes  $y$  to increase or decrease.
- The magnitude of coefficient is pretty meaningless.
- Coefficient  $\neq$  Marginal Effect
- $e^{b_2} \neq$  Marginal Effect. Almost as meaningless.
- A lot more math to figure out marginal effect.