Political Science 104

Lecture 16:
Dummy Variables and Interaction Terms

Nominal Level Variables as Independent Variables

- Recall that nominal level variables only tell us that observations fall into different unordered categories.
- Unlike ratio, interval, or ordinal variables, a "1-unit change" doesn't have any meaning.
- How can we use nominal level variables as independent variables in a regression?

Coding of Nominal Variables

- **Example**: A variable measuring gender:
  - Men = 1
  - Women = 2
- **Example**: A variable measuring race:
  - White = 1
  - Black = 2
  - Latino = 3
  - Asian = 4
  - Other = 5

Dummy Variables

- To use nominal level variables in a regression, we must recode them into a set of *dummy variables*.
- Dummy variables are coded 0 if an observation does not fall into a specific category, and 1 if it does fall into that category.
- After this recoding a 1-unit change will make sense -- it is the difference between being in and out of that category.

Examples of Dummy Variables

- Gender:
  - Women dummy variable (1 if woman, 0 if man)
- Race:
  - White dummy variable (1 if White, 0 otherwise)
  - Black dummy variable (1 if Black, 0 otherwise)
  - Latino dummy variable (1 if Latino, 0 otherwise)
  - Asian dummy variable (1 if Asian, 0 otherwise)
- One group in our data is the baseline category. All dummy variables measure the difference from the baseline category.

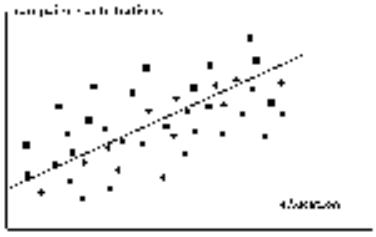
Dummy Variables in Regression

- Consider the regression:

  $$y = a + b*x + c*m + e$$

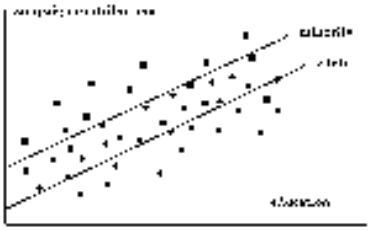
  where y = campaign contributions in $, x = years of education, and m is a dummy variable for minority (0 for white, 1 for minority).
- The regression line for whites: a + b*x
- The regression line for minorities: (a + c) + b*x
- The dummy variable adjusts the constant term.

Example: Regression Ignoring Differences Between Groups



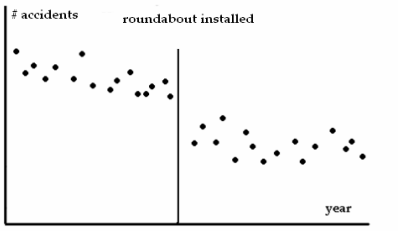
- circles = white, squares = minority

Example: Regression with Dummy Variable Distinguishing Groups



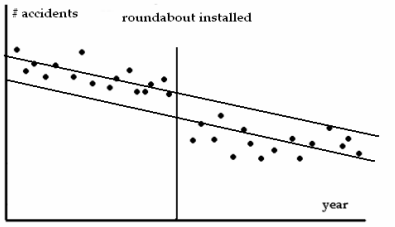
- Intercept for whites = a
- Intercept for minorities = a + c

Example: Dummy Variable in a Quasi-Experiment



- We want to test whether installing a traffic roundabout reduced traffic accidents.

Example: Dummy Variable in a Quasi-Experiment



- Specify a dummy variable separating the two time periods.

Example of Regression with Dummy Variables in SPSS

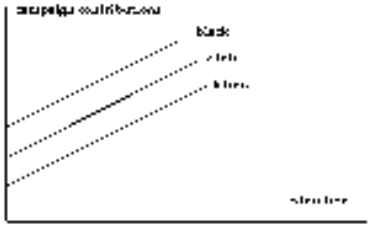
Dummy Variables in Regression, More than 2 Groups

- Consider the regression:

  y = a + b*x + c*m1 + d*m2 + e

  where y = campaign contributions in $, x = years of education, m1 is a dummy variable for black, and m2 is a dummy variable for latino.
- The regression line for white: a + b*x
- The regression line for black: (a + c) + b*x
- The regression line for latino: (a + d) + b*x.

Example: Regression with Dummy Variable Distinguishing 3 Groups



- Intercept for white = a, intercept for black = (a+c), intercept for latino = (a + d).

Dummy Variables in Regression, Two Dummy Variables

- Consider the regression:

  $$y = a + b*x + c*m + d*g + e$$

  where y = campaign contributions in $, x = years of education, m is a dummy variable for minority, and g is a dummy variable for women.
- The regression line for white male: a + b*x
- The regression line for minority male: (a + c) + b*x
- The regression line for white female: (a + d) + b*x.
- The regression line for minority female: (a + c + d) + b*x

Example: Regression with Two Dummy Variables

- Intercept for white male = a, intercept for minority male = (a+c), intercept for white female = (a + d), intercept for minority female = (a + c + d).