Stat 6021: Indicator Variables for Categorical Predictors

Read this after Section 1 in Guided Notes.

1 Number of Indicator Variables

The example below demonstrates why when we have a categorical variable with a classes, we will use $a-1$ indicator variables, each taking on the values of 0 and 1.

In a study of innovation in the insurance industry, an economist wishes to relate the speed with which an insurance innovation is adopted, y, to the size of the firm, x_1, and the type of firm, x_2. The variables are:

- y: number of months between the time the innovation was adopted in the industry and the time the given firm adopted the innovation.
- x_1: amount of total assets of the firm, which is quantitative.
- x_2: categorical with two classes (stock firm / mutual fund firm).

There are several ways of quantitatively identifying the classes of a categorical variable. **Indicator variables** that take on the values 0 and 1 are used widely. For example, since we have a categorical predictor with two classes, we may consider using two indicator variables, i.e.

$$\begin{aligned} I_1 &= \begin{cases} 1 & \text{if stock firm} \\ 0 & \text{otherwise;} \end{cases} \\ I_2 &= \begin{cases} 1 & \text{if mutual fund firm} \\ 0 & \text{otherwise;} \end{cases} \end{aligned}$$

and the model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 I_1 + \beta_3 I_2 + \epsilon.$$

Consider an example where $n = 4$, the first two companies are stock firms, and the last two companies are mutual fund firms. The design matrix, $\boldsymbol{X}$, becomes

$$\begin{bmatrix} 1 & x_{11} & 1 & 0 \\ 1 & x_{21} & 1 & 0 \\ 1 & x_{31} & 0 & 1 \\ 1 & x_{41} & 0 & 1 \end{bmatrix}$$

In this example, since the first column is exactly equal to the addition of column 3 and column 4, the matrix $\boldsymbol{X'X}$ is singular, i.e., the inverse of $\boldsymbol{X'X}$ does not exist, so we have no unique solution to the least-squares regression.

To get around this issue, we drop one of the indicator variables. In general, a categorical variable with a classes will be represented by $a-1$ indicator variables, each taking on the values of 0 and 1.

2 Interpreting Regression Coefficients with Indicator Variables: Additive Effect

Returning to our insurance example, we have the following model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 I_1 + \epsilon,$$

where x_1 = size of the firm and

$$I_1 = \begin{cases} 1 & \text{if stock firm} \\ 0 & \text{otherwise.} \end{cases}$$

Some call such a model a model with **additive effects** since the effect of each predictor is added into the model.

So the regression equations are:

Mutual fund firms: $\mathrm{E}\{Y\} = \beta_0 + \beta_1 x_1 + \beta_2(0) = \beta_0 + \beta_1 x_1$

Stock firms: $\mathrm{E}\{Y\} = \beta_0 + \beta_1 x_1 + \beta_2(1) = (\beta_0 + \beta_2) + \beta_1 x_1$

These equations inform us that the regression equations for both types of firms have the same slope and different intercepts, as shown in figure 1.

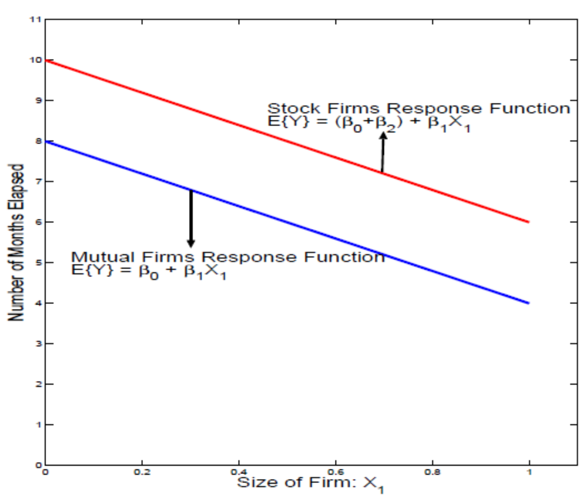

Figure 1: Regression Equations with Additive Effects

The coefficient β_2 for the indicator variable is the difference in the intercepts. Contextually, β_2 is the difference in the average response between stock firms and mutual fund firms, when comparing firms of similar size.

In general, β_2 shows how much higher (or lower) the mean response is for the class coded 1 than for the class coded 0, for any given level (or value) of x_1.

3 Interpreting Regression Coefficients with Indicator Variables: Interaction Effect Present

Additive effects assume that each predictor's effect on the response does not depend on the value(s) of the other predictor(s). **Interaction effects** allow the effect of one predictor on the response to depend on the value(s) of other predictor(s).

In this example, consider sale prices, y, of homes, with predictors x_1 = square footage of home and x_2 = whether home has air conditioning. We use the indicator variable $I_1 = 1$ if the home has air conditioning, and $I_1 = 0$ if the home does not have air conditioning.

Consider the following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 I_1 + \beta_3 x_1 I_1. \tag{1}$$

So the regression equations are:

No air conditioning: $E(Y) = \beta_0 + \beta_1 x_1$.

With air conditioning: $E(Y) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_1$.

We have different slopes and intercepts. The difference in the mean response between homes with and without airconditioning is no longer constant and depends on the square footage of the home. The difference in the mean response is given by $\beta_2 + \beta_3 x_1$.

In general, it is best to write out the regression equation for each class of the categorical variable.