Multiple Regression

p regression coefficients

$$y_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i \, ,$$ where $p-1$ is the # of predictors

$$= \sum_{k=0}^{p-1} \beta_k X_{ik} + \varepsilon_i \, ,$$ where $X_{i0} = 1$

Example: $p = 3 \Rightarrow 2$ predictors

$\beta_0 = 10$

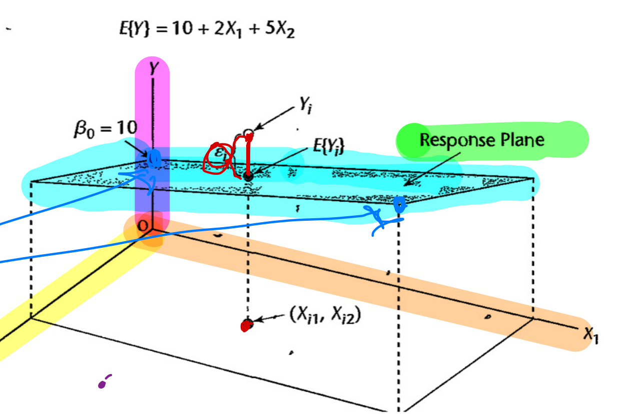
$\beta_1 = 2$

$\beta_2 = 5$

$$y_i = 10 + 2X_{i1} + 5X_{i2} + \varepsilon_i \, , \quad E\{\varepsilon_i\} = 0$$

$$E[y_i] = 10 + 2X_{i1} + 5X_{i2}$$

FIGURE 6.1 Response Function is a Plane—Sales Promotion Example.



these two points have different "heights" with respect to $E\{Y\}$ axis

Multiple Regression with Normal Errors

$$y_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i ,$$

$$= \sum_{k=0}^{p-1} \beta_k X_{ik} + \varepsilon_i , \quad \text{where } X_{i0} = 1$$

same model written different ways!

*since this sum starts at $k=0$, it's a sum over p terms

where
* $\beta_0, \beta_1, \dots, \beta_{p-1}$ parameters
* $X_{i1}, \dots, X_{ip-1}$ as fixed
* $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$ [This means $E\{\varepsilon_i\} = 0$, $\sigma^2\{\varepsilon_i\} = \sigma^2$]
* holds for $i = 1, \dots, n$

Regression Function: $E\{Y_i\} = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik}$

Model lets us conclude things about:
- Independence of Y_i, Y_j $i \neq j$, Y_i and Y_j are independent $i \neq j$
- Distribution of each Y_i $Y_i \sim N(\beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik}, \sigma^2)$
  * Variance of each Y_i, $\sigma^2\{Y_i\} = \sigma^2$

Multiple Linear Regression Model Uses:

$$y_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i$$

* multiple predictors
* qualitative predictor variables
    * binary predictor, e.g. biological sex

      could define corresponding predictor as

$$X_{i1} = \begin{cases} 1 & \text{if subject } i \text{ female} \\ 0 & \text{otherwise} \end{cases}$$

If X_{i1} is our predictor, we would have

$$y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i = \begin{cases} \beta_0 + \varepsilon_i & \text{if subject } i \text{ is male} \\ \beta_0 + \beta_1 + \varepsilon_i & \text{if subject } i \text{ is female} \end{cases}$$

Multiple Linear Regression Model Uses:

$$y_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i$$

* multiple predictors
* qualitative predictor variables
  * categorical predictor with more than two levels

example - y_i is student i's MCAS score

study was designed to include all students in Massachusetts in 2019, suppose we want to incorporate school as a predictor

suppose there are 10 schools in MA & our data

Define $X_{i1}, \ldots, X_{i9}$ where $X_{ik} = \begin{cases} 1 & \text{if student } i \text{ attended school } k \\ 0 & \text{otherwise} \end{cases}$

Regression function is:

$$\mathbb{E}[Y_i] = \begin{cases} \beta_0 & \text{if student } i \text{ attended school } 10 \\ \beta_0 + \beta_k & \text{if student } i \text{ attended school } k \end{cases}$$

Multiple Linear Regression Model Uses:

$$y_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i$$

Interpretation of β_k

change in mean response $E\{Y\}$ with a unit increase in predictor X_{ik} when all other predictors are held constant

* multiple predictors
* qualitative predictor variables
  - binary predictors
  - categorical predictors with more than two categories
* include polynomial terms when Y_i is not linear in covariate(s)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \varepsilon_i$$

If Y_i appears to be a quadratic function of X_{i1}

Define $X_{i2} = X_{i1}^2 \Rightarrow Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$

* interaction effects, e.g. Y_i is amount of time needed to produce lot i of parts

X_{i1} # of parts required for lot i

X_{i2} # of workers assigned to lot i

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} * X_{i2} + \varepsilon_i$$

, equivalent to defining $X_{i3} = X_{i1} * X_{i2}$

Matrix Notation for Multiple Linear Regression

$$\underset{\sim}{Y}_{n\times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \qquad X_{n\times p} = \underbrace{\left.\begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1\,p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2\,p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & & X_{n\,p-1} \end{bmatrix}\right\} n \text{ rows}}_{p \text{ columns}}$$

$$\underset{\sim}{\beta}_{p\times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} \qquad \underset{\sim}{\varepsilon}_{n\times 1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

Matrix Notation for Multiple Linear Regression

Let's write out

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

- $\underset{\sim}{\beta}$ is now a vector of parameters
- X is now a matrix of fixed constants
- $\underset{\sim}{\varepsilon}$ is now a random vector

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1\,p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2\,p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & & X_{n\,p-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$= \begin{bmatrix} \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_{p-1} X_{1\,p-1} + \varepsilon_1 \\ \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_{p-1} X_{2\,p-1} + \varepsilon_2 \\ \vdots \\ \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_{p-1} X_{n\,p-1} + \varepsilon_n \end{bmatrix}$$

Matrix Notation for Multiple Linear Regression

Let's write out

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

- $\underset{\sim}{\beta}$ is now a vector of parameters
- X is now a matrix of fixed constants
- $\underset{\sim}{\varepsilon}$ is now a random vector

$$E\{\underset{\sim}{\varepsilon}\} = \underset{\sim}{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\sigma^2\{\underset{\sim}{\varepsilon}\} = \sigma^2 I_n$$

diagonal $n \times n$ matrix with 1's on the diagonal, 0's everywhere else

$$= \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & 1 \end{bmatrix}$$