Interactions

y - response

x_1, x_2 - predictors

(1) $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

(2) $y = \gamma_0 + \gamma_1 x_1$

Train: $(\underset{\sim}{x}_1, y_1), \ldots (\underset{\sim}{x}_n, y_n)$

Q: When does $\hat{\gamma}_1 = \hat{\beta}_1$?

A: When $x_1 \perp x_2$.

$$\langle x_1, x_2 \rangle = x_1^T x_2 = \sum_{j=1}^{p} x_{1j} x_{2j}$$

$$= x_{11} x_{21} + x_{12} x_{22} + \ldots + x_{1p} x_{2p}$$

Recall

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Take $\beta_0 = 0$ (standardized data)

$$\underset{n\times 2}{X} = \begin{bmatrix} | & | \\ \underset{\sim}{x_1} & \underset{\sim}{x_2} \\ | & | \end{bmatrix}$$

$$X^T X = \begin{bmatrix} - x_1^T - \\ - x_2^T - \end{bmatrix} \begin{bmatrix} | & | \\ x_1 & x_2 \\ | & | \end{bmatrix}$$

$$= \begin{bmatrix} x_1^T x_1 & x_1^T x_2 \\ x_2^T x_1 & x_2^T x_2 \end{bmatrix}$$

x_1 $n\times 1$

x_1^T $1\times n$

$x_1^T x_1$ 1×1

If $x_1 \perp x_2$,

$$x_1^T x_2 = 0 = x_2^T x_1 = \langle x_1, x_2 \rangle$$

$$\hat{\beta} = \begin{pmatrix} x_1^T x_1 & 0 \\ 0 & x_2^T x_2 \end{pmatrix}^{-1} \begin{bmatrix} - x_1^T - \\ - x_2^T - \end{bmatrix} \begin{bmatrix} y \end{bmatrix}$$

$$\hat{\beta} = \begin{pmatrix} (X_1^T X_1)^{-1} & 0 \\ 0 & (X_2^T X_2)^{-1} \end{pmatrix} \begin{pmatrix} X_1^T y \\ X_2^T y \end{pmatrix}$$

$$= \begin{bmatrix} (X_1^T X_1)^{-1} X_1^T y \\ (X_2^T X_2)^{-1} X_2^T y \end{bmatrix} \begin{array}{c} \hat{\beta}_1 \\ \hat{\beta}_2 \end{array}$$

Interactions

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \underbrace{\beta_3 x_1 x_2}_{\text{interaction term}}$$

- linear in x_1
- — " — in x_2

(power is 1)

Case 1: x_2 is binary: 0 or 1

y income

x_1 education

x_2 gender (M = 0, F = 1)

Inc. Edu.

(1) $y = \beta_0 + \beta_1 X_1$

- Income does not depend on gender $(\beta_2 = 0)$

same edu. $\Rightarrow$ same predicted income

Inc. Edu gen.

(2) $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$

- Same education:

$$Inc_M - Inc_W = -\beta_2$$

- 1 yr extra of edu $\Rightarrow$ same extra income (β_1)

(3) $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$

$$\begin{array}{cccc} & \text{Edu } (x_1) & \text{gend } (x_2) & x_1x_2 \end{array}$$

$$X = \begin{bmatrix} 1 & 16 & 1 & 16 \\ 1 & 22 & 0 & 0 \\ 1 & 18 & 1 & 18 \\ 1 & 12 & 1 & 12 \\ 1 & 16 & 0 & 0 \\ \vdots & & & \end{bmatrix}$$

M: $X_2 = 0$

$$y = \beta_0 + \beta_1 X_1$$

W: $X_2 = 1$

$$y = \beta_0 + \beta_1 X_1 + \beta_2 + \beta_3 X_1$$
$$= (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_1$$

Equal: $\hat{\beta}_2, \hat{\beta}_3$ not signif.

Discrimination against women $\hat{\beta}_2 < 0, \hat{\beta}_3 < 0$

— " — men $\hat{\beta}_2 > 0, \hat{\beta}_3 > 0$

Case 2 x_2 categorical, 3 levels (>2)

y - income

x_1 - edu.

x_2 - race : Cauc., URM, Other

"C", "U", "O"

K levels → K-1 "dummy" variables

$$y = \text{"}\beta_0 + \beta_1 x_1 + \beta_2 x_2\text{"}$$

$$X = \begin{array}{cccc} & & C & U \\ \end{array}$$
$$X = \begin{bmatrix} 1 & 16 & 1 & 0 \\ 1 & 22 & 0 & 1 \\ 1 & 18 & 0 & 0 \\ 1 & 12 & 1 & 0 \\ 1 & 16 & 0 & 1 \\ & \vdots & & \end{bmatrix}$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_C + \beta_3 x_U$$

same edu. X_1

C: $\beta_0 + \beta_1 X_1 + \beta_2$

U: $\beta_0 + \beta_1 X_1 + \beta_3$

O: $\beta_0 + \beta_1 X_1$

baseline

β_2: diff. in income b/w C and O w/ same edu.

β_3: diff. b/w U and O w/ same edu.

Interaction model:

" $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \gamma X_1 X_2$ "

$\beta_0 + \beta_1 X_1 + \beta_2 X_G + \beta_3 X_U + \beta_4 X_1 X_C + \beta_5 X_1 U_U$

Slopes:

O: β_1

C: $\beta_1 + \beta_4$

U: $\beta_1 + \beta_5$