Example $Y_i = \beta_1 + \beta_2 X_i + u_i$

i	X_i	Y_i	$\bar{X}$	$\bar{Y}$	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$\hat{Y}_i$
1	8	1	5	3	$8-5=3$	$1-3=-2$	$7-\frac{4}{5}(8)=0.6$
2	4	3	5	3	$4-5=-1$	$3-3=0$	$7-\frac{4}{5}(4)=3.8$
3	6	2	5	3	$6-5=1$	$2-3=-1$	$7-\frac{4}{5}(6)=2.2$
4	2	6	5	3	$2-5=-3$	$6-3=3$	$7-\frac{4}{5}(2)=5.4$

OLS by hand

Use the OLS formulas

$$\hat{\beta}_2 = \frac{\sum (Y_i-\bar{Y})(X_i-\bar{X})}{\sum (X_i-\bar{X})^2} = \frac{-2\cdot 3 + 0\cdot(-1) + 1\cdot(-1) + 3\cdot(-3)}{3^2 + (-1)^2 + 1^2 + (-3)^2}$$

$$= \frac{-6-1-9}{9+1+1+9} = -\frac{4}{5}$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$= 3 - \left(-\frac{4}{5}\right)5 = 3 + \frac{4}{5}5 = 3+4$$

$$= 7$$

$$\underbrace{\hat{Y}_i = 7 - \frac{4}{5}X_i}_{\text{estimated regression line}}$$

Calculate R^2 based on the regression above.

$$R^2 = \frac{ESS}{TSS} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

$$R^2 = \frac{ESS}{TSS} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

$$= \frac{(0.6-3)^2 + (3.8-3)^2 + (2.2-3)^2 + (5.4-3)^2}{(1-3)^2 + (3-3)^2 + (2-3)^2 + (6-3)^2}$$

$$= \frac{12.8}{14}$$

$$= 0.91$$

The estimated model explains 91% of the variation in Y.

↳ model only includes one explanatory variable (X)

⇒ X "explains" 91% of the variation in Y.

Units of Measurement example

What happens to the slope estimate when we multiply Y_i by some constant λ?

Untransformed model: $Y_i = \beta_1 + \beta_2 X_i + u_i$

↳ estimated version: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

Transformation of Y: $Y_i^* = \lambda Y_i$

Transformed model: $Y_i^* = \beta_1^* + \beta_2^* X_i + u_i$

Slope estimator: $$\beta_2^* = \frac{\sum (X_i - \bar{X})(Y_i^* - \bar{Y}^*)}{\sum (X_i - \bar{X})^2}$$

$$= \frac{\sum (X_i - \bar{X})(\lambda Y_i - \lambda \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$= \lambda \hat{\beta}_2$$

Units of Measurement Practice problem

See slides for set-up.

Solution

Plug the transformed data into the intercept estimator:

$$\hat{\beta}_1 = \bar{\tilde{Y}} - \hat{\beta}_2 \bar{\tilde{X}}$$

$$= \frac{1}{n}\sum \tilde{Y}_i - \hat{\beta}_2 \frac{1}{n} \sum \tilde{X}_i$$

[by definition of the sample mean]

$$= \frac{1}{n}\sum (Y_i - \bar{Y}) - \hat{\beta}_2 \frac{1}{n} \sum (X_i - \bar{X})$$

[Plug in the transformed variable]

Aside: $\frac{1}{n}\sum(Y_i-\bar{Y})=0.$

Proof:

$$\begin{aligned}\frac{1}{n}\sum(Y_i-\bar{Y}) &= \frac{1}{n}\sum Y_i - \frac{1}{n}\sum\bar{Y} \\ &= \frac{1}{n}\sum Y_i - \frac{1}{n}n\bar{Y} \\ &= \underbrace{\frac{1}{n}\sum Y_i}_{\text{definition of } \bar{Y}} - \bar{Y} \\ &= \bar{Y}-\bar{Y} \\ &= 0\end{aligned}$$

Then $\hat{\beta}_1=(0)-\hat{\beta}_2(0)$

$=0$