1

STAT2008/STAT6038

Deriving the Least Squares equations for SLR

First, we need to remember some simple facts:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

so that

$$\sum_{i=1}^{n}(X_i - \overline{X}) = 0.$$

Similarly,

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

so that

$$\sum_{i=1}^{n}(Y_i - \overline{Y}) = 0.$$

Also,

$$\sum_{i=1}^{n} 1 = n.$$

2

Now, begin with the first of the equations to solve:

$$0 = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)$$

$$\Rightarrow 0 = \sum_{i=1}^{n} Y_i - \hat{\beta}_0 \sum_{i=1}^{n} 1 - \hat{\beta}_1 \sum_{i=1}^{n} X_i$$

$$\Rightarrow 0 = n\overline{Y} - n\hat{\beta}_0 - n\hat{\beta}_1 \overline{X}$$

$$\Rightarrow 0 = \overline{Y} - \hat{\beta}_0 - \hat{\beta}_1 \overline{X}$$

$$\Rightarrow \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}.$$

3

Now, moving to the second equation:

$$0 = \sum_{i=1}^{n} X_i(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)$$

$$\Rightarrow 0 = \sum_{i=1}^{n} X_i(Y_i - \{\overline{Y} - \hat{\beta}_1 \overline{X}\} - \hat{\beta}_1 X_i)$$

$$\Rightarrow 0 = \sum_{i=1}^{n} X_i((Y_i - \overline{Y}) - \hat{\beta}_1 (X_i - \overline{X}))$$

$$\Rightarrow 0 = \sum_{i=1}^{n} X_i(Y_i - \overline{Y}) - \hat{\beta}_1 \sum_{i=1}^{n} X_i(X_i - \overline{X})$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i(Y_i - \overline{Y})}{\sum_{i=1}^{n} X_i(X_i - \overline{X})}$$

4

---

there seems to be some things missing until you realise that

$$\sum_{i=1}^{n} \overline{X}(Y_i - \overline{Y}) = \overline{X}\sum_{i=1}^{n}(Y_i - \overline{Y}) = 0$$

and

$$\sum_{i=1}^{n} \overline{X}(X_i - \overline{X}) = \overline{X}\sum_{i=1}^{n}(X_i - \overline{X}) = 0,$$

so subtracting the former from the numerator in the expression for $\hat{\beta}_1$ and the latter from the denominator causes no change (since we are just subtracting zero from each of the numerator and denominator). Hence,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i(Y_i - \overline{Y})}{\sum_{i=1}^{n} X_i(X_i - \overline{X})}$$

$$= \frac{\sum_{i=1}^{n} X_i(Y_i - \overline{Y}) - \sum_{i=1}^{n} \overline{X}(Y_i - \overline{Y})}{\sum_{i=1}^{n} X_i(X_i - \overline{X}) - \sum_{i=1}^{n} \overline{X}(X_i - \overline{X})}$$

$$= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

$$= \frac{S_{xy}}{S_{xx}}.$$

5