JOHNS HOPKINS UNIVERSITY

Statistical Models and Regression

EN 625.661.82

Prepared by:

Sean Mahoney
sean.mahoney743@gmail.com
843-901-0071

Professor: Dr. Kelly Rooker

Title: Assignment 1
Date Submitted: 8/31/2018

Enclosed pages: 3

Sean Mahoney
EN: 625.661.82

Derive OLS estimator of B

$$y = Bx_i + e_i \ , \ E(e_i) = 0$$

$$S(B) = \sum (y_i - Bx_i)^2$$

$$0 = \partial S / \partial B_1 = 2 \sum_i (y_i - Bx_i) x_i \Rightarrow 0 = \sum x_i y_i - \sum Bx_i^2$$

$$\boxed{B = \sum x_i y_i / \sum x_i^2}$$

Derive estimator of variance of OLS estimator of B

$$B = \frac{\sum x_i y_i}{\sum x_i^2} \Rightarrow Var(B) = Var \frac{\sum x_i y_i}{\sum x_i^2} = Var\left(\frac{\sum x_i \cdot (Bx_i + e_i)}{\sum x_i^2}\right)$$

$$= \frac{\sum x_i^2}{(\sum x_i^2)^2} \cdot Var(Bx_i + e_i) = \boxed{\frac{\sigma^2}{\sum x_i^2}}$$

Derive the OLS estimators of B_0 and B_1

$$\hat{y} = B_1 x_1 + B_0 + \mathcal{E}_i$$

$$S(B_0, B_1) = \sum_i (y_i - \hat{y})^2 = \sum (y_i - (B_1 x_1 + B_0 + \mathcal{E}))^2 \ , \ E(\mathcal{E}) = 0$$

$$\partial S / \partial B_0 = 0 = -2\left(\sum y_i - \sum B_1 x_1 - \sum B_0\right)$$

$$\rightarrow nB_0 = n\bar{y} - B_1 n\bar{x} \Rightarrow \boxed{\hat{B}_0 = \bar{y} - \hat{B}_1 \bar{x}}$$

$$\partial S / \partial B_1 = 0 = -2 \sum (y_i - B_1 x_i - B_0) x_i = \sum x_i y_i - \sum B_1 x_i^2 - \sum B_0 x_i$$

$$0 = \sum x_i y_i - B_1 \sum x_i^2 - B_0 \sum x_i = \sum x_i y_i - B_1 \sum x_i^2 - (\bar{y} - \hat{B}_1 \bar{x}) \sum x_i$$

$$B_1 \sum x_i^2 - \hat{B}_1 \bar{x} \sum x_i = \sum x_i y_i - \bar{y} \sum x_i$$

$$B_1 \left(\sum x_i^2 - \bar{x} \sum x_i\right) = \sum x_i y_i - \bar{y} \sum x_i \Rightarrow \hat{B}_1 = \boxed{\frac{\sum x_i y_i - \frac{\sum y_i \sum x_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}}$$

$$\boxed{\begin{aligned}\tfrac{1}{n}\textstyle\sum x_i &= \bar{x}\\ n\bar{x} &= \textstyle\sum x_i\end{aligned}}$$

Derive Variance of OLS estimators:

First we need to use some equivalent functions:

$$\sum(x_i-\bar{x})(y_i-\bar{y}) = \sum(x_iy_i - x_i\bar{y} - y_i\bar{x} - \bar{x}\bar{y}) = \sum x_iy_i - \bar{y}\sum x_i - \bar{x}\sum y_i + n\bar{x}\bar{y}$$

$$= \sum x_iy_i - \bar{y}\bar{x}n - \bar{x}\bar{y}n + n\bar{x}\bar{y} = \sum x_iy_i - \bar{x}\bar{y}n = \sum x_iy_i - \bar{x}\sum y_i$$

also equal to $\boxed{\sum x_iy_i - \bar{y}\sum x_i}$

$$\hat{\beta}_1 = \frac{\sum x_iy_i - \frac{\sum x_i\sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} = \frac{\sum x_iy_i - \frac{n\bar{x}\cdot n\bar{y}}{n}}{\sum x_i^2 - \frac{(n\bar{x})^2}{n}} = \frac{\sum x_iy_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum x_i^2 - \sum\bar{x}^2}$$

$= \boxed{\dfrac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}}$ We will use this form to derive variance

$$Var\,\hat{\beta}_1 = Var\left(\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}\right), \quad y_i = \beta_0 + \beta_1x_i + e_i$$

$$\sum(x_i-\bar{x})(y_i-\bar{y}) = \sum y_i(x_i-\bar{x}) - \bar{y}\sum(x_i-\bar{x}), \quad \sum(x_i-\bar{x}) = n\bar{x}-n\bar{x} = 0$$

$= \sum y_i(x_i-\bar{x}) = \sum(x_i-\bar{x})(\beta_0+\beta_1x_i+e_i)$ → Now distribute + take variance

$Var(\,) = Var\left(\beta_0\sum(\cancelto{0}{x_i-\bar{x}}) + \beta_1\sum x_i(x_i-\bar{x}) + \sum(x_i-\bar{x})e_i\right)$

Since $\sum(x_i-\bar{x})=0$, This term becomes $Var\left(\beta_1(\bar{x}^2n(n-1)) + \sum(x_i-\bar{x})e_i\right)$

Variance of sums = Sum of variances → $\sum\left(Var(\beta_1\bar{x}^2n(n-1)) + Var[(x_i-\bar{x})e_i]\right)$

Variance of a constant = 0 ⇒ $\sum(x_i-\bar{x})^2\cdot Var(e_i)$, $Var(e_i)=\sigma^2$ (def)

Variance of numerator = $\sum(x_i-\bar{x})^2\cdot\sigma^2$, Denomitor is constant ⇒

$$Var = \frac{\sum(x_i-\bar{x})^2}{\left(\sum(x_i-\bar{x})^2\right)^2}\cdot\sigma^2 = \boxed{\frac{\sigma^2}{\sum(x_i-\bar{x})^2}} = Var(\hat{\beta}_1)$$

Variance of $\hat{\beta}_0$; $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$$Var(\hat{\beta}_0) = Var(\bar{y} - \hat{\beta}_1 \bar{x}) = Var(\bar{y}) - \bar{x}^2 Var\, \hat{\beta}_1 - 2\bar{x}\, Cov(\bar{y}, \hat{\beta}_1)$$

$$Var(\bar{y}) = \frac{\sigma^2}{n}, \quad Var(\hat{\beta}_1) = \bar{x}^2 \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}$$

$$Cov(\bar{y}, \hat{\beta}_1) = \sum \frac{(y_i - \bar{y})(\beta_i - \bar{\beta})}{N}$$, Since $\Sigma(y_i - \bar{y}) = 0$, Covariance is $\emptyset$

$$\Rightarrow Var(\hat{\beta}_0) = \frac{\sigma^2}{n} - \frac{\sigma^2 \bar{x}^2}{\Sigma(x_i - \bar{x})^2} = \boxed{\sigma^2 \left(\frac{1}{n} - \frac{\bar{x}^2}{\Sigma(x_i - \bar{x})^2} \right)}$$

$$S_{xx} = \Sigma(x_i - \bar{x})^2$$

Assumptions used for this assignment

- The error ε (or e_i) has expected value $= 0$ and constant variance of σ^2 and is normally distributed