2. Properties of estimators

① $\hat{\beta}_0, \hat{\beta}_1$ are both linear estimates

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})y_i - \sum_{i=1}^{n}(x_i-\bar{x})\bar{y}}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})y_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \sum_{i=1}^{n}\frac{(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}y_i = \sum_{i=1}^{n}c_i y_i$$

$$\bar{y}\sum_{i=1}^{n}(x_i-\bar{x}) = \bar{y}\left[\sum x_i - n\bar{x}\right] = 0$$

where $c_i = \dfrac{x_i-\bar{x}}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = \frac{\sum_{i=1}^{n}y_i}{n} - \bar{x}\sum_{i=1}^{n}c_i y_i = \sum_{i=1}^{n}\frac{1}{n}y_i - \sum_{i=1}^{n}c_i\bar{x}y_i = \sum_{i=1}^{n}\left(\frac{1}{n}-c_i\bar{x}\right)y_i = \sum_{i=1}^{n}d_i y_i$$

where $d_i = \frac{1}{n} - c_i\bar{x}$

fitted value: $\hat{y}_j = \hat{\beta}_0 + \hat{\beta}_1 x_j$

$$= \sum_{i=1}^{n}d_i y_i + \sum_{i=1}^{n}c_i y_i \cdot x_j$$

$$= \sum_{i=1}^{n}(d_i + c_i x_j)y_i$$

② When $x_0 = \bar{x}$, then $\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$

$$= \bar{y} - \hat{\beta}_1\bar{x} + \hat{\beta}_1\bar{x} = \bar{y}$$

$\Rightarrow$ the linear regression line passes through $(\bar{x}, \bar{y})$

③ unbiaseness

$$\sum_{i=1}^{n}c_i = \frac{\sum_{i=1}^{n}(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = 0$$

$$\sum_{i=1}^{n}c_i x_i = \frac{\sum_{i=1}^{n}x_i(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 + \bar{x}\sum_{i=1}^{n}(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})}$$

$$= 1 + 0 = 1$$

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} \left(\frac{1}{n} - c_i \bar{x}\right)$$

$$= 1 - \bar{x} \sum_{i=1}^{n} c_i = 1$$

$$\sum_{i=1}^{n} (d_i x_i) = \sum_{i=1}^{n} \left(\frac{1}{n} - c_i \bar{x}\right) x_i$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i - \bar{x} \sum_{i=1}^{n} c_i x_i$$

$$= \bar{x} - \bar{x} \cdot 1 = 0$$

$$E(\hat{\beta}_1) = E\left(\sum_{i=1}^{n} c_i y_i\right) = E\left(\sum_{i=1}^{n} c_i (\beta_0 + \beta_1 x_i + \varepsilon_i)\right)$$

$$= E\left(\beta_0 \sum_{i=1}^{n} c_i + \beta_1 \sum_{i=1}^{n} c_i x_i + \sum_{i=1}^{n} c_i \varepsilon_i\right)$$

$$= E(0 + \beta_1 + 0)$$

$$= \beta_1$$ as estimator equals expectation it is unbiased

$$E(\hat{\beta}_0) = E\left(\sum_{i=1}^{n} d_i y_i\right) = E\left[\sum_{i=1}^{n} d_i (\beta_0 + \beta_1 x_i + \varepsilon_i)\right]$$

$$= E\left[\sum_{i=1}^{n} d_i \beta_0 + \sum_{i=1}^{n} d_i x_i \beta_1\right]$$

$$= E[\beta_0] = \beta_0$$

④ Variance

$$Var(\hat{\beta}_1) = Var\left(\sum_{i=1}^{n} c_i y_i\right)$$

$$= \sum_{i=1}^{n} c_i^2 \, Var(y_i)$$

$$= \sum_{i=1}^{n} \left[\frac{x_i - \bar{x}}{\sum_{i=1}^{n} (x_i - \bar{x})^2}\right]^2 \cdot \sigma^2$$

$$= \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{S_{xx}^2} \cdot \sigma^2$$

$$= \frac{\sigma^2}{S_{xx}}$$

$$Var(\hat{\beta}_0) = Var\left(\sum_{i=1}^{n} d_i y_i\right)$$

$$= \sum_{i=1}^{n} d_i^2 \, Var(y_i)$$

$$= \sum_{i=1}^{n} \left(\frac{1}{n} - c_i \bar{x}\right)^2 \cdot \sigma^2$$

$$= \sum_{i=1}^{n} \left(\frac{1}{n^2} - \frac{2c_i \bar{x}}{n} + c_i^2 \bar{x}^2\right) \cdot \sigma^2$$

$$= \left(\frac{1}{n} - 0 + \frac{\bar{x}^2}{S_{xx}}\right) \cdot \sigma^2$$

$$= \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)$$

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

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

$$= Cov\left(\frac{1}{n}\sum_{i=1}^{n} y_i, \sum_{i=1}^{n} c_i y_i\right) - \bar{x}\, Var(\hat{\beta}_1)$$

$Cov(y_i, y_j) = 0$
$i \neq j$

$$= \sigma^2 \left(\frac{1}{n}\sum_{i=1}^{n} c_i\right) - \bar{x}\, Var(\hat{\beta}_1)$$

$$= -\bar{x}\, Var(\hat{\beta}_1)$$

$$= -\bar{x} \frac{\sigma^2}{S_{xx}}$$

$$Var(\hat{y}_i) = Var(\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$$= Var(\hat{\beta}_0) + Var(\hat{\beta}_1 x_i) + 2Cov(\hat{\beta}_0, \hat{\beta}_1 x_i)$$

$$= \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right) + \frac{x_i^2 \sigma^2}{S_{xx}} - 2x_i \frac{\bar{x}}{S_{xx}} \sigma^2$$

$$= \sigma^2 \left(\frac{1}{n} + \frac{(\bar{x} - x_i)^2}{S_{xx}}\right)$$

⑤ $RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$

$$E(RSS) = E\left[\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2\right]$$
$$= E\left[\sum_{i=1}^{n} (y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i)^2\right]$$
$$= E\left[\sum_{i=1}^{n} \left[(y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x})\right]^2\right]$$
$$= E\left[\sum_{i=1}^{n} \left[(y_i - \bar{y})^2 - 2\hat{\beta}_1 (x_i - \bar{x})(y_i - \bar{y}) + \hat{\beta}_1^2 (x_i - \bar{x})^2\right]\right]$$
$$= \sum_{i=1}^{n} \left[E(y_i - \bar{y})^2\right] - 2E\left[\hat{\beta}_1 S_{xy}\right] + E\left[\hat{\beta}_1^2 S_{xx}\right]$$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

$$= \sum_{i=1}^{n} \left[E(y_i - \bar{y})^2\right] - 2 \cdot E\left(\frac{S_{xy}^2}{S_{xx}}\right) + E\left[\frac{S_{xy}^2}{S_{xx}}\right]$$
$$= \sum_{i=1}^{n} \left[E(y_i - \bar{y})^2\right] - E\left(\frac{S_{xy}}{S_{xx}}\right)$$

1) $E\left(\frac{S_{xy}^2}{S_{xx}}\right) = E(\hat{\beta}_1^2 \cdot S_{xx})$
$$= S_{xx} E(\hat{\beta}_1^2)$$
$$= S_{xx}\left[E(\hat{\beta}_1)^2 + Var(\hat{\beta}_1)\right]$$
$$= S_{xx}\left[\beta_1^2 + \frac{\sigma^2}{S_{xx}}\right]$$
$$= S_{xx}\beta_1^2 + \sigma^2$$

2) $\sum_{i=1}^{n} \left[E(y_i - \bar{y})^2\right] = \sum_{i=1}^{n} \left[(E(y_i - \bar{y}))^2 + var(y_i - \bar{y})\right]$
$$= \sum_{i=1}^{n} \left\{\left[(\beta_0 + \beta_1 x_i) - (\beta_0 + \beta_1 \bar{x})\right]^2 + (n-1)\frac{\sigma^2}{n^2} + \frac{(n-1)^2}{n^2}\sigma^2\right\}$$
$$= \sum_{i=1}^{n} \beta_1^2 (x_i - \bar{x})^2 + \frac{n-1}{n}\sigma^2 + \frac{(n-1)^2}{n}\sigma^2$$

$$Var(y_i - \bar{y}) = Var\left(-\frac{y_1}{n} - \frac{y_2}{n} - \dots - \frac{n-1}{n} y_i - \dots\right)$$
$$= Var\left(\frac{-y_1}{n}\right) + var\left(\frac{-y_2}{n}\right) + \dots + var\left(\frac{n-1}{n} y_i\right) + \dots + var\left(\frac{-y_n}{n}\right)$$
$$= (n-1)\frac{\sigma^2}{n^2} + \frac{(n-1)^2}{n^2}\sigma^2$$

02/03/2022

$$E(RSS) = \sum_{i=1}^{n} \beta_1^2 (x_i - \bar{x})^2 + \frac{n-1}{n}\sigma^2 + \frac{(n-1)^2}{n}\sigma^2 - (S_{xx}\beta_1^2 + \sigma^2)$$

$$= \sigma^2 \left(\frac{(n-1)^2 + n-1}{n} - \frac{n}{n} \right)$$

Error somewhere...

$$= \sigma^2 (n-2)$$

(maybe fixed tomorrow)