Supplementary Document for Simple Linear Regression

Calculation of Expectation of SS_{Res}

Note that

$$\widehat{y_i} = \widehat{\beta}_0 + \widehat{\beta}_1 x_i = \bar{y} + \widehat{\beta}_1 (x_i - \bar{x}) .$$

As,

$$SS_{Res} = \sum_{i=1}^{n} (y_i - \widehat{y_i})^2 = \sum_{i=1}^{n} \left(y_i^2 + \widehat{y_i^2} - 2y_i\widehat{y_i}\right) ,$$

the expectation of SS_{Res} can be written as

$$E\left(SS_{Res}\right) = \sum_{i=1}^{n} \left[E(y_i^2) + E(\widehat{y_i}) - 2E(y_i\widehat{y_i})\right] .$$

Now, we will calculate each expectation in the previous expression.

$$\begin{aligned} E(y_i^2) &= \text{Var}\,(y_i) + E^2\,(y_i) = \sigma^2 + (\beta_0 + \beta_1 x_i)^2 \\ E(\widehat{y_i^2}) &= \text{Var}\,(\widehat{y_i}) + E^2\,(\widehat{y_i}) \\ &= \text{Var}(\bar{y}) + (x_i - \bar{x})^2\,\text{Var}\left(\widehat{\beta}_1\right) + 2\,(x_i - \bar{x})\,\text{Cov}\left(\bar{y}, \widehat{\beta}_1\right) + (\beta_0 + \beta_1 x_i)^2 . \end{aligned}$$

$$\text{Cov}\left(\bar{y}, \widehat{\beta}_1\right) = \text{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} y_i, \sum_{j=1}^{n} \frac{x_j - \bar{x}}{S_{xx}} y_j\right) = \frac{\sigma^2}{S_{xx}} \sum_{j=1}^{n} (x_j - \bar{x}) = 0.$$

Thus,

$$E(\widehat{y_i^2}) = \frac{\sigma^2}{n} + \frac{(x_i - \bar{x})^2\,\sigma^2}{s_{xa}} + (\beta_0 + \beta_i x_i)^2 .$$

Again,

$$E(y_i\widehat{y_i}) = \text{Cov}\,(y_i,\, \widehat{y_i}) + E(y_i)E(\widehat{y_i}).$$

Now,

$$\text{Cov}\,(y_i,\, \widehat{y_i}) = \text{Cov}\,(y_i,\, \bar{y}) + (x_i - \bar{x})\text{Cov}\left(y_i,\, \widehat{\beta}_1\right) = \frac{\sigma^2}{n} + \frac{\sigma^2\,(x_i - \bar{x})^2}{S_{xx}},$$

implies

$$E\left(y_i\widehat{y}_i\right)=\frac{\sigma^2}{n}+\left(\beta_0+\beta_1x_i\right)^2+\frac{\left(x_i-\bar{x}\right)^2\sigma^2}{S_{xx}}.$$

Therefore,

$$\begin{aligned}E\left(SS_{R_2S}\right)&=\sum_{i=1}^{n}\Bigg[\sigma^2+\left(\beta_0+\beta_1x_i\right)^2+\frac{\sigma^2}{n}+\frac{\left(x_i-\overline{x}\right)\sigma^2}{S_{xx}}+\left(\beta_0+\beta_1x_i\right)^2\\&\qquad-2\left(\beta_0+\beta_1x_i\right)^2-\frac{2\sigma^2}{n}-\frac{2\sigma^2(x_i-\bar{x})^2}{S_{xx}}\Bigg]\\&=(n-2)\sigma^2.\end{aligned}$$

Calculation for Expectation of SS_{Reg}

$$SS_{Reg}=\sum_{i=1}^{n}\left(\widehat{y}_i-\bar{y}\right)^2=\sum_{i=1}^{n}\left(\widehat{y}_i^2+\bar{y}^2-2\widehat{y}_i\bar{y}\right).$$

Now,

$$E(\bar{y}^2)=\mathrm{Var}(\bar{y})+E^2(\bar{y})=\frac{\sigma^2}{n}+\left(\beta_0+\beta_1\bar{x}\right)^2.$$

$$E\left(\widehat{y}_i\bar{y}\right)=E\left[\bar{y}^2+\left(x_i-\bar{x}\right)\bar{y}\widehat{\beta}_1\right]=\frac{\sigma^2}{n}+\left(\beta_0+\beta_1\bar{x}\right)^2+\left(x_i-\bar{x}\right)E\left(\bar{y}\widehat{\beta}_1\right).$$

Moreover,

$$\mathrm{Cov}\left(\bar{y},\widehat{\beta}_1\right)=\mathrm{Cov}\left(\frac{1}{n}\sum_{i=1}^{n}y_i,\sum_{j=1}^{n}\frac{x_j-\bar{x}}{S_{xx}}y_j\right)=\sigma^2\sum_{i=1}^{n}\frac{x_i-\bar{x}}{nS_{xx}}=0.$$

Thus,

$$E\left(\bar{y}\widehat{\beta}_1\right)=\mathrm{Cov}\left(\bar{y},\widehat{\beta}_1\right)+E(\bar{y})E\left(\widehat{\beta}_1\right)=\left(\beta_0+\beta_1\bar{x}\right)\beta_1$$

Therefore,

$$E\left(\widehat{y}_i\bar{y}\right)=\frac{\sigma^2}{n}+\left(\beta_0+\beta_1\bar{x}\right)^2+\left(x_i-\bar{x}\right)\beta_1\left(\beta_0+\beta_1\bar{x}\right).$$

Combining all these expectations, we get

$$\begin{aligned}E\left(SS_{Reg}\right)&=\sum_{i=1}^{n}\Bigg[\frac{\sigma^2}{n}+\frac{\left(x_i-\bar{x}\right)^2\sigma^2}{S_{xx}}+\left(\beta_0+\beta_1x_i\right)^2+\frac{\sigma^2}{n}+\left(\beta_0+\beta_1\bar{x}\right)^2\\&\qquad-2\frac{\sigma^2}{n}-2\left(\beta_0+\beta\bar{x}\right)^2-2\beta_1\left(\beta_0+\beta_1\bar{x}\right)\left(x_i-\bar{x}\right)\Bigg]\\&=\sigma^2+\sum_{i=1}^{n}\left[\left(\beta_0+\beta_1x_i\right)^2-\left(\beta_0+\beta_1\bar{x}\right)^2-2\left(\beta_0+\beta_1\bar{x}\right)\left(\beta_0+\beta_1x_i-\beta_0-\beta_1\bar{x}\right)\right]\\&=\sigma^2+\beta_1^2S_{xx}.\end{aligned}$$

Calculation of Expectation of SS_T

$$SS_T = SS_{Reg} + SS_{Res}.$$

Therefore,

$$E\left(SS_{Reg}\right) = E\left(SS_{Reg}\right) + E\left(SS_{Res}\right) = (n-1)\sigma^2 + \beta_1^2 S_{xx}.$$