

Proof from Week 2, slide 28

With grateful acknowledgment to Becky Lin, we start by taking the expected value of the RSS summation,

$$\begin{aligned}\mathbb{E} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 &= \sum_{i=1}^n \mathbb{E}(Y_i - \hat{Y}_i)^2 \\&= \sum_{i=1}^n \text{var}(Y_i - \hat{Y}_i) \quad \text{because } \beta_0 \text{ and } \beta_1 \text{ are unbiased} \\&= \sum_{i=1}^n \text{var}(Y_i - \hat{Y} - b_1(x_i - \bar{x})) \quad \text{from e.g. slide 18, Week 2} \\&= \sum_{i=1}^n [\text{var}(Y_i - \bar{Y}) - 2\text{cov}((Y_i - \bar{Y}), b_1(x_i - \bar{x})) + (x_i - \bar{x})^2 \text{var}(b_1)] \\&= \sum_{i=1}^n \left[\text{var}(e_i - \bar{e}_i) - 2\text{cov}((Y_i - \bar{Y})(x_i - \bar{x}), b_1) + (x_i - \bar{x})^2 \frac{\sigma^2}{S_{xx}} \right] \\&= (n-1)\sigma^2 - 2\text{cov}\left(\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y}), b_1\right) + \sigma^2\end{aligned}$$

Using the equation on slide 19, Week 2:

$$\begin{aligned}\mathbb{E} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 &= (n-1)\sigma^2 - 2\text{cov}(b_1 S_{xx}, b_1) + \sigma^2 \\ &= (n-1)\sigma^2 - S_{xx} 2\text{var}(b_1) + \sigma^2 \\ &= (n-1)\sigma^2 - S_{xx} \frac{2\sigma^2}{S_{xx}} + \sigma^2 \\ &= (n-2)\sigma^2\end{aligned}$$

Therefore, S^2 on slide 28 is an unbiased estimate of σ^2