REMARKS & EXTRA EXERCISES

Chapter 3

1. BLUE versus MLE

There are two paramters to be estimated in the simple linear regression (SLR) model. For these, there are two methods of estimation are favored and widely used: one is the least square estimation (LSE) and the other is the maximum likelihood estimation (MLE). In addition to this, the mean square of the errors is also an important component to be estimated in SLR model.

From Exercise 3.21(b) on p. 128, we know that LSEs of β_0, β_1 are coincided with MLEs of β_0, β_1. Then, are the least square estimate (LSE) of σ^2 and the maximum likelihood estimate (MLE) of σ^2 the same? If they are not the same, which one is better to use?

(a) The maximum likelihood estimmator of σ^2 :

The likelihood function of Y_i is

$$L\left(\beta_0, \beta_1, \sigma^2; \mathbf{y}\right) = \prod_{i=1}^{n} f_Y\left(y_i\right) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right)^2\right\},$$

the log-likelihood is

$$\log L\left(\beta_0, \beta_1, \sigma^2; \mathbf{y}\right) = -\frac{n}{2}\log\left(2\pi\right) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right)^2.$$

To get the MLE. we set

$$\frac{\partial \log L\left(\beta_0, \beta_1, \sigma^2; \mathbf{y}\right)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right)^2 = 0.$$

Solving equation for σ^2 gives us the MLE

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2.$$

(b) The least squares estimates (LSE) of σ^2 is the mean square error.

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-2}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2.$$

We see that the MLE $\tilde{\sigma}^2$ is smaller than the LSE $\hat{\sigma}^2$ of σ^2.

Example 3.1: (Ex. #3.21, p. 128) From the data, we calculate that the LSE $\hat{\sigma}^2 = 7.4176$ and the MLE $\tilde{\sigma}^2 = 5.5632$, which is smaller than the MLE. We note that the LSE is BLUE and the MLE is not unbised.

Theorem 3.1: *In SLR model,* $E\left(s^2\right) = \sigma^2$.

Proof: The proof is left as an exercise. □

Theorem 3.2: *The maximum likelihood estimator $\tilde{\sigma}^2$ is asmptotically unbised for σ^2.*

Proof: The proof is left as an exercise. □