Problem Set 1: Unbiased Estimator of the Error Variance

This assignment will guide you through the derivations needed to determine what is a unbiased estimator of the error variance in the context of a univariate linear regression. This assignment may be quite challenging. Good Luck!

Consider the univariate linear regression model

$$y_t = \alpha + \beta x_t + u_t, \quad t = 1, \ldots, T. \tag{1}$$

where the regressors are non-stochastic (fixed) and the disturbances have zero mean and are uncorrelated and homoscedastic with variance equal to σ^2, i.e., $\mathbb{E}[u_t] = 0$, $\mathbb{C}[u_t, u_s] = 0$, $t \neq s$, and $\mathbb{V}[u_t] = \mathbb{E}[u_t^2] = \sigma^2$. The aim of this problem set will be to derive that

$$s^2 = \frac{1}{T-2} \sum_{t=1}^{T} \hat{u}_t^2$$

is an unbiased estimator of the disturbance variance σ^2, where $\hat{u}_t = y_t - \hat{y}_t$, $\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t$, and $\hat{\alpha}$ and $\hat{\beta}$ are the OLS estimators of α and β, respectively. The OLS estimates $\hat{\alpha}$ and $\hat{\beta}$ are given by

$$\hat{\beta} = \frac{\sum_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2}, \quad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x},$$

which can also be expressed as

$$\hat{\beta} = \sum_{t=1}^{T} w_t y_t, \quad \hat{\alpha} = \sum_{t=1}^{T} q_t y_t,$$

where

$$w_t = \frac{x_t - \bar{x}}{\sum_{t=1}^{T} (x_t - \bar{x})^2}, \quad q_t = \frac{1}{T} - \bar{x} \cdot w_t.$$

Question 1 - Verify simple properties.

Verify the following 7 properties:

1. $\sum_{t=1}^{T} w_t = 0$
2. $\sum_{t=1}^{T} w_t x_t = 1$
3. $\sum_{t=1}^{T} q_t = 1$
4. $\sum_{t=1}^{T} q_t x_t = 0$
5. $\sum_{t=1}^{T} w_t^2 = \frac{1}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$
6. $\sum_{t=1}^{T} q_t^2 = \frac{\frac{1}{T}\sum_{t=1}^{T} x_t^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$
7. $\sum_{t=1}^{T} q_t w_t = \frac{-\bar{x}}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$

As you work through the problem set, you should always look back at these 8 properties and see which one can help you in each step.

For example, with properties 1 and 2 in hand, it is easy to show (as we did in class) that:

$$\hat{\beta} = \sum_{t=1}^{T} w_t y_t = \sum_{t=1}^{T} w_t \left(\alpha + \beta x_t + u_t\right) = \beta + \sum_{t=1}^{T} w_t u_t.$$

Likewise, using properties 3 and 4 it is easy to show (as we did in class) that:

$$\hat{\alpha} = \sum_{t=1}^{T} q_t y_t = \sum_{t=1}^{T} q_t \left(\alpha + \beta x_t + u_t\right) = \alpha + \sum_{t=1}^{T} q_t u_t.$$

Question 2 - Computing the expression for the sum of the residuals.

Show that the residuals from the regression described by Eq. 1 can be expressed as:

$$\hat{u}_t = -\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] + u_t.$$

And, therefore,

$$\hat{u}_t^2 = \left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right]^2 + u_t^2 - 2\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] u_t$$

The sum of the residuals therefore can be expressed as:

$$\sum_{t=1}^{T} \hat{u}_t^2 = \sum_{t=1}^{T} \left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right]^2 + \sum_{t=1}^{T} u_t^2 - \sum_{t=1}^{T} 2\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] u_t$$

The expected sum of the residuals therefore can be expressed as:

$$\sum_{t=1}^{T} \mathbb{E}\left[\hat{u}_t^2\right] = \sum_{t=1}^{T} \mathbb{E}\left[\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right]^2\right] + \sum_{t=1}^{T} \mathbb{E}\left[u_t^2\right] - \sum_{t=1}^{T} \mathbb{E}\left[2\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] u_t\right]$$

The expression above can be broken down into 3 components:

- Component 1: $\sum_{t=1}^{T} \mathbb{E}\left[2\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] u_t\right]$
- Component 2: $\sum_{t=1}^{T} \mathbb{E}\left[\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right]^2\right]$
- Component 3: $\sum_{t=1}^{T} \mathbb{E}\left[u_t^2\right]$

The next 3 questions will make you work on each of these components.

Question 3 - Working with Component 1.

Note that:

$$(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t = \left(\sum_{s=1}^{T} q_s u_s\right) + \left(\sum_{s=1}^{T} w_s u_s\right) x_t.$$

Using the expression above, show that:

$$\mathbb{E}\left[\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] u_t\right] = \left(q_t + w_t x_t\right)\sigma^2,$$

and, therefore, that

$$\sum_{t=1}^{T} \mathbb{E}\left[2\left[(\hat{\alpha}-\alpha) + \left(\hat{\beta}-\beta\right) x_t\right] u_t\right] = 4\sigma^2.$$

Question 4 - Working with Component 2

Show that

$$\mathbb{E}\left[\left[(\hat{\alpha}-\alpha)+\left(\hat{\beta}-\beta\right)x_t\right]^2\right]=\sigma^2\left[\frac{\frac{1}{T}\sum_{t=1}^{T}x_t^2}{\sum_{t=1}^{T}\left(x_t-\bar{x}\right)^2}+\frac{x_t^2}{\sum_{t=1}^{T}\left(x_t-\bar{x}\right)^2}-2\frac{\bar{x}\cdot x_t}{\sum_{t=1}^{T}\left(x_t-\bar{x}\right)^2}\right],$$

and, therefore, that

$$\sum_{t=1}^{T}\mathbb{E}\left[\left[(\hat{\alpha}-\alpha)+\left(\hat{\beta}-\beta\right)x_t\right]^2\right]=2\sigma^2.$$

(This can be quite challenging)

Question 5 - Working with Component 3

Show that

$$\sum_{t=1}^{T}\mathbb{E}\left[u_t^2\right]=T\sigma^2$$

(This should be very easy)

Final Step

This final step is easy and I will solve it for you. Once we have obtained the expressions for the three components in questions 3, 4, and 5 we simply have to put all the pieces together to obtain an unbiased estimator of the error variance. To be precise:

$$\sum_{t=1}^{T}\hat{u}_t^2=\sum_{t=1}^{T}\mathbb{E}\left[(\hat{\alpha}-\alpha)+\left(\hat{\beta}-\beta\right)x_t\right]^2+\sum_{t=1}^{T}\mathbb{E}\left[u_t^2\right]-\sum_{t=1}^{T}\mathbb{E}\left[2\left[(\hat{\alpha}-\alpha)+\left(\hat{\beta}-\beta\right)x_t\right]u_t\right]$$

simplifies to

$$\sum_{t=1}^{T}\mathbb{E}\left[\hat{u}_t^2\right]=2\cdot\sigma^2+T\cdot\sigma^2-4\cdot\sigma^2=(T-2)\cdot\sigma^2.$$

Divide by $(T-2)$ on both sides the expression above to obtain the unbiased estimator of the error variance:

$$\mathbb{E}\left[\frac{1}{T-2}\sum_{t=1}^{T}\hat{u}_t^2\right]=\sigma^2.$$