Jared Yu

Module 2 Discussion

1. A typical simple linear regression model is $y = \beta_0 + \beta_1 x + \varepsilon$, where y is a response variable (also often called dependent variable), x is an independent variable (also often called regressor), and ε is a random error with mean (also called expectation) zero. Thus y and ε are random variables. The regressor x is either a random variable or a non-random (also often called fixed) variable. A set of n indepedendent paired data $(y_1, x_2), \cdots, (y_n, x_n)$ follow this model. Before the n paired data values are available, we construct the ordinary least squares (OLS) estimator for β_0 and β_1 as described in Chapter 2 of the Textbook.
    a. Discuss whether the assumption of "the constant variance σ^2" (see (2.1) in the Textbook) is required for the construction of the OLS estimators.

Ans:

On p. 12 of the Textbook, it says that the errors are assumed to have mean zero and unknown variance σ^2. In equation (2.2b) of the Textbook, it shows the following formula:

$$Var(y|x) = Var(\beta_0 + \beta_1 x + \varepsilon) = \sigma^2. \quad (1)$$

This says that the variance of y does not depend on x. Also, when the errors are uncorrelated, that also implies that the responses y_i's are uncorrelated. This would mean then that the data is homoscedastic with constant variance. A counterexample would be a dataset where rather than the paired data points following a straight channel, they would fan out towards the end in something called heteroscedasticity. The effect then is that the variance of y is dependent on x. Following this then would be that the error terms, ε_i's, also have an increasing variance that is dependent on x.

The Textbook calls it the *method of least squares*, but for this discussion it will be equivalent to OLS. When using OLS to construct estimators for the parameters β_0 and β_1, it is solving for the objective function:

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 . \quad (2)$$

This step does not necessarily require that there is constant variance σ^2. It is merely solving for $\hat{\beta}_0$ and $\hat{\beta}_1$. However, the property of constant variance is important for something like the *Gauss-Markov theorem*, where it states that the OLS model will lead to the *best linear unbiased estimators (B.L.U.E.)*, as long as the assumptions of $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$, and that the errors are uncorrelated are held (p. 19). In other words, without constant variance, the OLS estimates do not have the property of B.L.U.E. In such a case, it would make more sense to apply a different model that accounts for heteroscedasticity.

    b. Discuss whether the constant variance assumption in (2.1) of the Textbook is required for unbiasedness of OLS estimators.

Ans:

To show that the OLS estimators are unbiased, the following two equations must hold true:

$$E(\hat{\beta}_0) = \beta_0 \text{ and } E(\hat{\beta}_1) = \beta_1. \quad (3)$$

The formula for $\hat{\beta}_0$ and $\hat{\beta}_1$ are as follows:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \text{ and } \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}}. \quad (4)$$

Furthermore, the formula for $\hat{\beta}_1$ can be further rewritten as follows:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \sum_{i=1}^{n} c_i y_i \,, \tag{5}$$

where $c_i = \frac{(x_i - \bar{x})}{S_{xx}}$. To show the unbiasedness, we will first start with $\hat{\beta}_1$ (pp. 18-19):

$$E(\hat{\beta}_1) = E\left(\sum_{i=1}^{n} c_i y_i\right) \tag{6}$$

In equation (6), we have simply applied the expectation to $\hat{\beta}_1$.

$$= \sum_{i=1}^{n} c_i E(y_i) \tag{7}$$

In equation (7), we move the expectation into the summation, based on the *linearity of expectation*. Furthermore, we are treating c_i as a constant, fixed term, since it consists entirely of x. It is important to note that this step is not possible if x is also a random variable.

$$= \sum_{i}^{n} c_i(\beta_0 + \beta_1 x_i) = \beta_0 \sum_{i=1}^{n} c_i + \beta_1 \sum_{i=1}^{n} c_i x_i \tag{8}$$

In equation (8), the $E(y_i)$ is rewritten as $\beta_0 + \beta_1 x_i$, since $E(\varepsilon_i) = 0$. Then the summation is distributed to each part of this new term.

$$\sum_{i=1}^{n} c_i = \sum_{i=1}^{n} \frac{(x_i - \bar{x})}{\left[\sum_{j=1}^{n}(x_j - \bar{x})^2\right]} = \frac{1}{\left[\sum_{j=1}^{n}(x_j - \bar{x})^2\right]}\left(\sum_{i=1}^{n} x_i - n\bar{x}\right) = 0 \tag{9}$$

In equation (9), it is shown that the $\sum_{i=1}^{n} c_i$ term equals 0, since $\sum_{i=1}^{n} x_i = n\bar{x}$.

$$\sum_{i=1}^{n} c_i x_i = \sum_{i=1}^{n} \frac{(x_i - \bar{x})x_i}{\left[\sum_{j=1}^{n}(x_j - \bar{x})^2\right]} = \frac{1}{\left[\sum_{j=1}^{n}(x_j - \bar{x})^2\right]} \sum_{i=1}^{n} (x_i - \bar{x})x_i \tag{10}$$

In equation (10), the formula $\sum_{i=1}^{n} c_i x_i$ has been expanded. To show that it is equivalent to 1, it must be shown that the numerator $\sum_{i=1}^{n}(x_i - \bar{x})x_i$ is equivalent to the denominator $\sum_{j=1}^{n}(x_j - \bar{x})^2$. This will be shown as follows:

$$\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x}) = \sum_{i=1}^{n}(x_i - \bar{x})x_i - \bar{x}\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n}(x_i - \bar{x})x_i - 0. \tag{11}$$

So, it follows that $\sum_{i=1}^{n} c_i = 0$ and $\sum_{i=1}^{n} c_i x_i = 1$, therefore equation (8) evaluates to β_1. Thus far it has been shown then that $\hat{\beta}_1$ is an unbiased estimator.

Lastly, we will show that $\hat{\beta}_0$ is unbiased.

$$E(\hat{\beta}_0) = E(\bar{y} - \hat{\beta}_1 \bar{x}) \tag{12}$$

In equation (12), the formula for $\hat{\beta}_0$ has simply been placed inside the expectation. This formula lacks a β_0, to obtain it, we must expand $\bar{y}$. This can be done as follows:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i + \varepsilon_i) = \beta_0 + \beta_1 \bar{x} + \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i = \beta_0 + \beta_1 \bar{x} \tag{13}$$

In equation (13), the formula for $\bar{y}$ has been to show β_0 along with other terms. An important property also is that $\sum_{i=1}^{n} \varepsilon_i = 0$, which can be shown as follows:

$$\sum_{i=1}^{n} \varepsilon_i = \sum_{i=1}^{n} (y_i - \hat{y}_i) = \sum_{i=1}^{n} \left[y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i \right) \right] \tag{14}$$

In equation (14), the formula for ε_i has been expanded and the $\hat{y}_i$ term also expanded.

$$= \sum_{i=1}^{n} \left\{ y_i - \left[\left(\bar{y} - \hat{\beta}_1 \bar{x} \right) + \hat{\beta}_1 x_i \right] \right\} = \sum_{i=1}^{n} \left(y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 \bar{x} \right) \tag{15}$$

In equation (15), the $\hat{\beta}_0$ term has been expanded and the " $-$ " sign distributed inside the brackets.

$$= \sum_{i=1}^{n} (y_i - \bar{y}) = \sum_{i=1}^{n} y_i - n\bar{y} = 0 \tag{16}$$

In equation (16), the $\hat{\beta}_1 \bar{x}$ terms cancel out, leaving the familiar $\sum_{i=1}^{n} (y_i - \bar{y})$ which sums to 0.

$$\cdots = E\left(\bar{y} - \hat{\beta}_1 \bar{x} \right) = E\left[(\beta_0 + \beta_1 \bar{x}) - \hat{\beta}_1 \bar{x} \right] = E\left[\beta_0 + \bar{x} \left(\beta_1 - \hat{\beta}_1 \right) \right] \tag{17}$$

In equation (17), the result from equation (13) is used to replace $\bar{y}$ in equation (12).

$$= E(\beta_0) + \bar{x} E\left(\beta_1 - \hat{\beta}_1 \right) = \beta_0 \tag{18}$$

In equation (18), the $E\left(\beta_1 - \hat{\beta}_1 \right) = 0$ since it has been shown already that $\hat{\beta}_1$ is unbiased. Also, again it is important to note that it is possible to pull out $\bar{x}$ since it is being considered fixed.

The above steps required only that we are able to construct the OLS estimators. The steps taken did not require that any sort of uncorrelation exists amongst the y_i terms. *Therefore, constant variance is not required for the unbiasedness of OLS estimators.*