

EC 306: Introductory Econometrics

Class Problem Sheet 1: Understanding What Makes a Good Estimator in Simple Contexts

Sketches of Answers are Provided in Below After Each Question

1. Let Y_i for $i = 1, \dots, N$ be a random sample from a probability distribution with mean μ and variance 1 (i.e. $E(Y_i) = \mu$ and $var(Y_i) = 1$ for $i = 1, \dots, N$). Interest centres on estimating μ and four different possible estimators are considered:

$$\hat{\mu}_1 = \frac{\sum_{i=1}^N Y_i}{N}$$

$$\hat{\mu}_2 = \frac{Y_1 + Y_N}{2}$$

$$\hat{\mu}_3 = \frac{1}{2} \left[\frac{\sum_{i=1}^{N_1} Y_i}{N_1} + \frac{\sum_{i=N_1+1}^N Y_i}{N - N_1} \right] \text{ where } 1 < N_1 < N$$

$$\hat{\mu}_4 = \frac{\sum_{i=1}^N Y_i}{N - 1}$$

i) Which of these is an unbiased estimator for μ ?

Sketch answer: I provide the derivations in detail for one of the estimators, $\hat{\mu}_4$, the derivations for the others are similar.

An estimator is a random variable (since it depends on the random sample of data) and has a mean. An unbiased estimator is one where this mean is equal to the true value (intuitively, an estimate will never be exactly the same as the true value. However, an unbiased estimator produces estimates which are, on average, true). Thus, we want to know if

$$E(\hat{\mu}_4) = \mu.$$

Here are the steps required:

$$\begin{aligned}
E(\hat{\mu}_4) &= E\left(\frac{\sum_{i=1}^N Y_i}{N-1}\right) \\
&= \frac{1}{N-1} E\left(\sum_{i=1}^N Y_i\right) \\
&= \frac{\sum_{i=1}^N E(Y_i)}{N-1} \\
&= \frac{\sum_{i=1}^N \mu}{N-1} \\
&= \frac{N\mu}{N-1} \\
&\neq \mu.
\end{aligned}$$

Therefore, this estimator is biased. Now let us go through the equal signs in the equation above. The first equal sign just uses the definition of $\hat{\mu}_4$. The second uses the fact that a constant can be taken outside an expected value. The third uses the result that "the expected value of a sum of independent random variables equals the sum of their expected values". The fourth equal sign uses the fact that $E(Y_i) = \mu$, while the fifth is just a property of sums (i.e. adding up N μ 's gives you $N\mu$).

Using similar methods, you can show $\hat{\mu}_1$ (the sample mean) is unbiased, $\hat{\mu}_2$ (which simply averages the first and last observations) is also unbiased. $\hat{\mu}_3$ (which takes one part of the data and averages it, then takes the rest of the data and averages it and then averages these averages) is also unbiased.

ii) The term "asymptotically unbiased" means that an estimator becomes unbiased as N goes to ∞ . Which of these is an asymptotically unbiased estimator of μ ?

Sketch answer: All of the unbiased estimators are also asymptotically unbiased. For instance, in part i) you showed $E(\hat{\mu}_1) = \mu$. There is no N in this formula – this unbiasedness property holds for any sample size. Thus we only need to look more deeply at $\hat{\mu}_4$. Previously, we derived $E(\hat{\mu}_4) = \frac{N\mu}{N-1}$. In your previous study of maths, you have probably discussed limits. Using such methods it can easily be shown that $\frac{N}{N-1}$ goes to 1 as N goes to ∞ and, thus, even though $\hat{\mu}_4$ is biased, it is asymptotically unbiased. If you have not studied limits before, just try putting larger and larger values for N in a calculator and you will see $\frac{N}{N-1}$ gets closer and closer to one as N gets bigger (e.g. if $N = 3$, then $\frac{N}{N-1} = 1.5$, if $N = 100$ then $\frac{N}{N-1} = 1.01$, if $N = 10,000$ then $\frac{N}{N-1} = 1.0001$, etc.).

iii) Derive the variance of each of the estimators that you found to be unbiased in part i).

Sketch answer: I provide the derivations in detail for one of the estimators, $\hat{\mu}_4$, the derivations for the others are similar. Here are the steps required:

$$\begin{aligned}
\text{var}(\hat{\mu}_4) &= \text{var}\left(\frac{\sum_{i=1}^N Y_i}{N-1}\right) \\
&= \left(\frac{1}{N-1}\right)^2 \text{var}\left(\sum_{i=1}^N Y_i\right) \\
&= \frac{\sum_{i=1}^N \text{var}(Y_i)}{(N-1)^2} \\
&= \frac{\sum_{i=1}^N 1}{(N-1)^2} \\
&= \frac{N}{(N-1)^2}.
\end{aligned}$$

Now let us go through the equal signs in the equation above. The first equal sign just uses the definition of $\hat{\mu}_4$. The second uses the fact that when a constant is taken outside a variance, it must be squared. The third uses the result that "the variance of a sum of independent random variables equals the sum of their variances". The fourth equal sign uses the fact that $\text{var}(Y_i) = 1$, while the fifth is just a property of sums (i.e. adding up N ones gives you N).

Using similar methods you can show $\text{var}(\hat{\mu}_1) = \frac{1}{N}$, $\text{var}(\hat{\mu}_2) = \frac{1}{2}$, $\text{var}(\hat{\mu}_3) = \frac{1}{4} \left(\frac{1}{N_1} + \frac{1}{N-N_1} \right)$.

iv) Use the concept of "efficiency" and your results from part iii) to discuss which estimator you think is best.

Sketch answer: Why am I asking these questions (apart from giving you practice in working with the expected value and variance operators)? I want you to see how that, in any econometric problem, there are always lots of possible estimators and even (as we have seen in part i) lots of possible unbiased estimators. How do we choose between all these possibilities? Efficiency is one way of doing so. It says "among all possible unbiased estimators, choose the one with the smallest variance". If we compare the variances of our unbiased estimators (these variances were calculated in part ii), you can (with a bit of manipulation), see that $\text{var}(\hat{\mu}_1)$ is the smallest. Thus, $\hat{\mu}_1$ is more efficient than the other estimators.

v) Show that the setup in this question is equivalent to the linear regression model with only an intercept (or equivalently, the simple regression model discussed in lectures with $X_i = 1$ for $i = 1, \dots, N$) where μ plays the role of the intercept (and the error variance in the regression model is set to 1).

Sketch answer: In a regression model with just an intercept, it can be seen that the expected value of y is β and the variance of y is just the error variance (in this question this is 1). So this regression model implies y is a random sample with a constant mean and a variance of 1. But this is the setup of this question.

vi) Using part v), what does the Gauss Markov theorem say is the best linear unbiased estimator (BLUE) for μ ?

Sketch answer: For this setup, the Gauss Markov theorem says that OLS is BLUE. But if you write out what the OLS estimator is in this case (i.e. when $X_i = 1$), you find:

$$\hat{\beta} = \frac{\sum X_i y_i}{\sum X_i^2} = \frac{\sum y_i}{N}.$$

But this is just the sample mean. So the sample mean is the best linear unbiased estimator (in this question we have labelled the sample mean as $\hat{\mu}_1$).

2. The simple linear regression model without intercept was discussed in the lectures and is given by:

$$Y_i = \beta X_i + \varepsilon_i,$$

where X_i is a scalar. Assume the classical assumptions hold.

Note: These questions are similar to part i) and to material done in lecture, so I will not provide detailed answers.

i) (Review of lecture material). The ordinary least squares estimator is given by:

$$\hat{\beta} = \frac{\sum X_i Y_i}{\sum X_i^2}.$$

Show that $E(\hat{\beta}) = \beta$ (and, thus, that $\hat{\beta}$ is unbiased) and $\text{var}(\hat{\beta}) = \frac{\sigma^2}{\sum X_i^2}$

This derivation was done in lectures

ii) Consider another estimator for β :

$$\tilde{\beta} = \frac{\sum Y_i}{\sum X_i}.$$

Is this an unbiased estimator for β ?

Answer: Yes it is unbiased.

iii) Calculate $\text{var}(\tilde{\beta})$.

iv) Use the concept of "efficiency" and your results from the previous parts of the question to discuss which estimator you think is best.

Answer: The key idea here is that $\hat{\beta}$ and $\tilde{\beta}$ are both unbiased estimators, so to choose one over another you have to look at their variances. With a little work you can show $\text{var}(\hat{\beta}) < \text{var}(\tilde{\beta})$ and, thus, that the OLS estimator is more efficient than the other estimator introduced in part ii). However, you do not really have to do the derivations in part iii) and iv). You can simply say: "The Gauss Markov theorem tells me that the OLS estimator has the smallest variance (among all linear unbiased estimators), therefore I know automatically that it must have a smaller variance than $\tilde{\beta}$ ".