

Sketches of Answers to the Questions in Chapter 3
Sketches of Answers are Provided in Below After Each Question

1. This exercise offers some practice at using statistical tables. See Appendix B, Definitions B.10 through B.13 for definitions of distributions and more examples.

a) Suppose $X \sim N(-10, 16)$. Calculate $\Pr(X \leq -14)$.

b) Suppose you have a test statistic, X , which under H_0 , has a Chi-square distribution with 20 degrees of freedom. In your data set, the test statistic is calculated to be 30. Do you reject H_0 at the 5% level of significance? Do you reject H_0 at the 10% level of significance?

c) Suppose you have a test statistic, X , which under a certain hypothesis: H_0 , has a t_{10} distribution. Using your data set, the test statistic is calculated to be 3.0. Do you reject H_0 at the 5% level of significance? Do you reject H_0 at the 1% level of significance?

d) Suppose you have a test statistic, X , which under a certain hypothesis: H_0 , has an $F_{2,60}$ distribution. In your data set, the test statistic is calculated to be 20. Do you reject H_0 at the 5% level of significance?

Answer: I will not provide answers here. Very similar derivations are provided in Appendix B where the various distributions used in this question are defined.

2. 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 five different estimators are considered:

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

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

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

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

$$\hat{\mu}_5 = Y_1 + \frac{\sum_{i=2}^N Y_i}{N - 1}$$

a) Which of these are unbiased estimators for μ ?

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 these two averages) is also unbiased. $\hat{\mu}_4$ is biased.

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

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 $\hat{\mu}_4$ and $\hat{\mu}_5$. 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.).

$\hat{\mu}_5$ is asymptotically biased.

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

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

d) Use the concept of efficiency and your results from part iii) to discuss which estimator is best.

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.

e) 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 this chapter 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).

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.

f) Using part e), what does the Gauss Markov theorem say is the best linear unbiased estimator for μ ?

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 $\hat{\mu}_1$).

3. The simple linear regression model is given by:

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

Assume that the classical assumptions hold.

Note: These questions are similar to 2 and to material done in textbook, so I will not provide detailed answers.

a) Consider an estimator for β :

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

Is this an unbiased estimator for β ?

Answer: Yes this is an unbiased estimator.

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

c) Compare $\tilde{\beta}$ to the OLS estimator $\hat{\beta}$. Use the concept of efficiency and your results from the previous parts of the question to discuss which estimator you think is best. Hint: To simplify the derivations, you may use the following result:

$$\frac{1}{\sum X_i^2} < \frac{N}{(\sum X_i)^2}.$$

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 $\tilde{\beta}$ ".

4. In the body of the chapter, we used the classical assumptions to derive the maximum likelihood estimator for β , assuming that σ^2 was known. This question asks you to extend these results to the case where σ^2 is unknown.

a) What is the likelihood function, $L(\beta, \sigma^2)$ which treats both β and σ^2 as being unknown?

Answer: This derivation is exactly the same as in the chapter, except for the interpretation (i.e. σ^2 is now treated as being unknown)

b) Show that the maximum likelihood estimator of β is still $\hat{\beta}$, the OLS estimator and that the maximum likelihood estimator for σ^2 , which we denote by $\hat{\sigma}^2$ is:

$$\hat{\sigma}^2 = \frac{\sum \hat{\varepsilon}_i^2}{N},$$

where $\hat{\varepsilon}_i$ for $i = 1, \dots, N$ are the OLS residuals.

Answer: The derivation that maximum likelihood estimator of β is $\hat{\beta}$, is exactly the same as in the chapter. Deriving the maximum likelihood estimator for σ^2 is a straightforward calculus question (take first derivatives of the log likelihood with respect to σ^2 , set to zero and then solve). I will not provide details.

c) Using the fact that the OLS estimator of σ^2 (i.e. s^2) is unbiased to show that $\hat{\sigma}^2$ is a biased estimator of σ^2 . Show that the bias of $\hat{\sigma}^2$ disappears as $N \rightarrow \infty$. Note that this implies $\hat{\sigma}^2$ is asymptotically unbiased (see Exercise 2 for a definition of this term).

Answer: Since

$$\hat{\sigma}^2 = \frac{N-1}{N} s^2,$$

we can take expected values of both sides and use the fact that s^2 is unbiased to derive a formula for $E(\hat{\sigma}^2)$ as follows:

$$\begin{aligned} E(\hat{\sigma}^2) &= E\left(\frac{N-1}{N} s^2\right) \\ &= \frac{N-1}{N} E(s^2) \\ &= \frac{N-1}{N} \sigma^2 \end{aligned}$$

Looking at this expression, you can see that $\hat{\sigma}^2$ is biased, but asymptotically unbiased (see question 2, part ii) for help with the proof of this)

5. Let Y_i for $i = 1, \dots, N$ be a random sample from a Normal probability distribution with mean μ and variance 1 (i.e. $E(Y_i) = \mu$ and $var(Y_i) = 1$ for $i = 1, \dots, N$). Exercise 2 provides five different estimators for μ . Derive a 95% confidence interval for μ for each of these estimators. Can the Gauss-Markov theorem be used to tell you which of these confidence intervals is narrower than the others?

Answer: I will not provide a complete answer here since the form of the derivation is exactly the same as in the section on "Deriving a confidence interval" in the chapter. That derivation began with

$$\hat{\beta} \text{ is } N \left(\beta, \frac{\sigma^2}{\sum X_i^2} \right)$$

Here you just have to replace this with the correct estimator and its mean and variance and go through exactly the same steps to derive the confidence interval. E.g. using results from exercise 2,

$$\hat{\mu}_4 \text{ is } N \left(\frac{N\mu}{N-1}, \frac{N}{(N-1)^2} \right)$$

You can use this to make a probability statement about $\hat{\mu}_4$ which can then be re-arranged to isolate μ in the middle of an inequality to produce the confidence interval.

The Gauss-Markov theorem can be used to say that, among the unbiased estimators, $\hat{\mu}_1$ will have the smallest variance. Looking at the formula for the confidence interval, you can see that this means that $\hat{\mu}_1$ will have narrower confidence intervals than any other unbiased estimator. However, the Gauss Markov theorem can say nothing about the biased estimators (it relates only to unbiased estimators).