Econ 103 Review

Jake Anderson

ECON 104, UCLA

Fall 2026

Outline

1. Question 1: One-Sided t-Test
2. Question 2: Signing Omitted Variable Bias
3. Question 3: Auxiliary Regressions and the VIF
4. Question 4: Testing a Linear Combination

Question 1: Labor Supply

The theory of labor supply says people offer more hours of work at higher wages. Let HRSWK be usual hours worked per week and WAGE the hourly wage, with

$$\text{HRSWK} = \beta_1 + \beta_2 \,\text{WAGE} + e.$$

From a sample of $N = 9{,}799$ individuals, the fitted regression is $\widehat{\text{HRSWK}} = 41.58 + 0.011\,\text{WAGE}$, and the estimated variances and covariance of the least squares estimators are

	Intercept	WAGE
Intercept	0.02324	−0.00067
WAGE	−0.00067	0.00003

The degrees of freedom are $N - 2 = 9{,}797$, so critical values may be read from the ∞ row of the t table. Test, at the 5% level, whether the data support the theory of labor supply, and compute the p-value for your test. State a clear conclusion.

Source: Econ 103 Final, Summer 2026 Session A, Question 23

Question 1, Step 1: Write the Hypotheses

What does the theory say? People work more hours at higher wages.

Question 1, Step 1: Write the Hypotheses

What does the theory say? People work more hours at higher wages.
Which parameter measures this? β_2, the change in weekly hours for a one-dollar increase in the hourly wage.

Question 1, Step 1: Write the Hypotheses

What does the theory say? People work more hours at higher wages.
Which parameter measures this? β_2, the change in weekly hours for a one-dollar increase in the hourly wage.
What does the theory say about β_2? More hours at higher wages means $\beta_2 > 0$.

Question 1, Step 1: Write the Hypotheses

What does the theory say? People work more hours at higher wages.
Which parameter measures this? β_2, the change in weekly hours for a one-dollar increase in the hourly wage.
What does the theory say about β_2? More hours at higher wages means $\beta_2 > 0$.
Where does the theory go? The claim we want the data to support goes in the *alternative* hypothesis H_1. The null H_0 is everything else:

$$H_0 : \beta_2 \leq 0 \qquad \text{vs} \qquad H_1 : \beta_2 > 0$$

Question 1, Step 1: Write the Hypotheses

What does the theory say? People work more hours at higher wages.
Which parameter measures this? β_2, the change in weekly hours for a one-dollar increase in the hourly wage.
What does the theory say about β_2? More hours at higher wages means $\beta_2 > 0$.
Where does the theory go? The claim we want the data to support goes in the *alternative* hypothesis H_1. The null H_0 is everything else:

$$H_0 : \beta_2 \leq 0 \qquad \text{vs} \qquad H_1 : \beta_2 > 0$$

Which tail? H_1 points to values of β_2 *greater* than 0 $\implies$ this is a **right-tail test**. We reject H_0 only if the test statistic is large and positive.

Question 1, Step 2: Find the Variance of b_2

What is in the table? Diagonal entries are variances. Off-diagonal entries are covariances.

	Intercept	WAGE
Intercept	$\widehat{\text{Var}}(b_1) = 0.02324$	$\widehat{\text{Cov}}(b_1, b_2) = -0.00067$
WAGE	$\widehat{\text{Cov}}(b_1, b_2) = -0.00067$	$\mathbf{\widehat{Var}(b_2) = 0.00003}$

Question 1, Step 2: Find the Variance of b_2

What is in the table? Diagonal entries are variances. Off-diagonal entries are covariances.

	Intercept	WAGE
Intercept	$\widehat{\text{Var}}(b_1) = 0.02324$	$\widehat{\text{Cov}}(b_1, b_2) = -0.00067$
WAGE	$\widehat{\text{Cov}}(b_1, b_2) = -0.00067$	$\mathbf{\widehat{Var}(b_2) = 0.00003}$

Which entry do we need? The test is about β_2, the coefficient on WAGE. Its variance is in the WAGE row and the WAGE column:

$$\widehat{\text{Var}}(b_2) = 0.00003$$

Question 1, Step 3: Compute the Standard Error of b_2

The standard error is the square root of the variance:

$$\text{se}(b_2) = \sqrt{\widehat{\text{Var}}(b_2)}$$

Question 1, Step 3: Compute the Standard Error of b_2

The standard error is the square root of the variance:

$$\text{se}(b_2) = \sqrt{\widehat{\text{Var}}(b_2)}$$

Plug in:

$$\text{se}(b_2) = \sqrt{0.00003}$$

Question 1, Step 3: Compute the Standard Error of b_2

The standard error is the square root of the variance:

$$\text{se}(b_2) = \sqrt{\widehat{\text{Var}}(b_2)}$$

Plug in:

$$\text{se}(b_2) = \sqrt{0.00003}$$

Compute:

$$\sqrt{0.00003} = 0.005477$$

Question 1, Step 3: Compute the Standard Error of b_2

The standard error is the square root of the variance:

$$\text{se}(b_2) = \sqrt{\widehat{\text{Var}}(b_2)}$$

Plug in:

$$\text{se}(b_2) = \sqrt{0.00003}$$

Compute:

$$\sqrt{0.00003} = 0.005477$$

Check: square the answer and confirm we get the variance back:

$$0.005477 \times 0.005477 = 0.0000300 \quad \checkmark$$

Question 1, Step 4: Compute the t-Statistic

Formula: the t-statistic compares the estimate to the value of β_2 at the boundary of H_0. The boundary of H_0: $\beta_2 \leq 0$ is $\beta_2 = 0$:

$$t = \frac{b_2 - 0}{\text{se}(b_2)}$$

Question 1, Step 4: Compute the t-Statistic

Formula: the t-statistic compares the estimate to the value of β_2 at the boundary of H_0. The boundary of H_0: $\beta_2 \leq 0$ is $\beta_2 = 0$:

$$t = \frac{b_2 - 0}{\text{se}(b_2)}$$

Plug in $b_2 = 0.011$ (the WAGE coefficient in the fitted regression) and $\text{se}(b_2) = 0.005477$ (Step 3):

$$t = \frac{0.011 - 0}{0.005477} = \frac{0.011}{0.005477}$$

Question 1, Step 4: Compute the t-Statistic

Formula: the t-statistic compares the estimate to the value of β_2 at the boundary of H_0. The boundary of H_0: $\beta_2 \leq 0$ is $\beta_2 = 0$:

$$t = \frac{b_2 - 0}{\text{se}(b_2)}$$

Plug in $b_2 = 0.011$ (the WAGE coefficient in the fitted regression) and $\text{se}(b_2) = 0.005477$ (Step 3):

$$t = \frac{0.011 - 0}{0.005477} = \frac{0.011}{0.005477}$$

Compute:

$$t = 2.01$$

Question 1, Step 4: Compute the t-Statistic

Formula: the t-statistic compares the estimate to the value of β_2 at the boundary of H_0. The boundary of H_0: $\beta_2 \leq 0$ is $\beta_2 = 0$:

$$t = \frac{b_2 - 0}{\text{se}(b_2)}$$

Plug in $b_2 = 0.011$ (the WAGE coefficient in the fitted regression) and $\text{se}(b_2) = 0.005477$ (Step 3):

$$t = \frac{0.011 - 0}{0.005477} = \frac{0.011}{0.005477}$$

Compute:

$$t = 2.01$$

In words: the estimate 0.011 is 2.01 standard errors above 0.

Question 1, Step 5: Find the Critical Value

Three pieces of information pick the critical value:

1. Significance level: $\alpha = 0.05$
2. Number of tails: one (right tail, from Step 1) $\implies$ all 5% of the rejection region is in the right tail
3. Degrees of freedom: 9,797 $\implies$ use the ∞ row of the t table

Question 1, Step 5: Find the Critical Value

Three pieces of information pick the critical value:

1. Significance level: $\alpha = 0.05$
2. Number of tails: one (right tail, from Step 1) $\implies$ all 5% of the rejection region is in the right tail
3. Degrees of freedom: 9,797 $\implies$ use the ∞ row of the t table

Read the table: we need the value with 95% of the distribution below it and 5% above it:

$$t_c = t_{(0.95,\ \infty)} = 1.645$$

Question 1, Step 5: Find the Critical Value

Three pieces of information pick the critical value:

1. Significance level: $\alpha = 0.05$
2. Number of tails: one (right tail, from Step 1) $\implies$ all 5% of the rejection region is in the right tail
3. Degrees of freedom: 9,797 $\implies$ use the ∞ row of the t table

Read the table: we need the value with 95% of the distribution below it and 5% above it:

$$t_c = t_{(0.95,\ \infty)} = 1.645$$

Rejection rule: reject H_0 if $t > 1.645$.

Question 1, Step 6: Make the Decision

Compare the test statistic (Step 4) to the critical value (Step 5):

$$t = 2.01 \qquad \text{vs} \qquad t_c = 1.645$$

Question 1, Step 6: Make the Decision

Compare the test statistic (Step 4) to the critical value (Step 5):

$$t = 2.01 \qquad \text{vs} \qquad t_c = 1.645$$

$$2.01 > 1.645$$

Question 1, Step 6: Make the Decision

Compare the test statistic (Step 4) to the critical value (Step 5):

$$t = 2.01 \qquad \text{vs} \qquad t_c = 1.645$$

$$2.01 > 1.645$$

The test statistic is in the rejection region $\implies$ reject $H_0 : \beta_2 \leq 0$.

Question 1, Step 7: Compute the p-Value

Definition: for a right-tail test, the p-value is the probability of a test statistic at least as large as the one we observed, if H_0 is true:

$$p = P(t > 2.01)$$

Question 1, Step 7: Compute the p-Value

Definition: for a right-tail test, the p-value is the probability of a test statistic at least as large as the one we observed, if H_0 is true:

$$p = P(t > 2.01)$$

Distribution: with 9,797 degrees of freedom, the t distribution is the standard normal Z:

$$p = P(Z > 2.01)$$

Question 1, Step 7: Compute the p-Value

Definition: for a right-tail test, the p-value is the probability of a test statistic at least as large as the one we observed, if H_0 is true:

$$p = P(t > 2.01)$$

Distribution: with 9,797 degrees of freedom, the t distribution is the standard normal Z:

$$p = P(Z > 2.01)$$

Use the normal table: the table gives the area *below* a value, $\Phi(2.01) = 0.9778$. The area *above* is 1 minus that:

$$p = 1 - \Phi(2.01) = 1 - 0.9778 = 0.0222$$

Question 1, Step 7: Compute the p-Value

Definition: for a right-tail test, the p-value is the probability of a test statistic at least as large as the one we observed, if H_0 is true:

$$p = P(t > 2.01)$$

Distribution: with 9,797 degrees of freedom, the t distribution is the standard normal Z:

$$p = P(Z > 2.01)$$

Use the normal table: the table gives the area *below* a value, $\Phi(2.01) = 0.9778$. The area *above* is 1 minus that:

$$p = 1 - \Phi(2.01) = 1 - 0.9778 = 0.0222$$

Compare to α: $0.0222 < 0.05 \implies$ reject H_0. This is the same decision as Step 6.

Question 1, Step 8: Conclusion

Conclusion

- **Decision:** Reject $H_0: \beta_2 \leq 0$ in favor of $H_1: \beta_2 > 0$ at the 5% level.
- **Evidence:** $t = 2.01 > 1.645$, and $p = 0.0222 < 0.05$.
- **Estimate in words:** a \$1 increase in the hourly wage is associated with 0.011 more hours worked per week. (A \$10 increase is associated with 0.11 more hours per week, about 7 minutes.)
- **Answer to the question:** the data support the theory of labor supply. Hours worked increase with the wage.

Outline

1. Question 1: One-Sided t-Test
2. Question 2: Signing Omitted Variable Bias
3. Question 3: Auxiliary Regressions and the VIF
4. Question 4: Testing a Linear Combination

Question 2: Two Omitted Variables

Two researchers each leave a variable out of their model.

I. A housing economist regresses Price on SQFT alone, omitting LotSize. Larger lots command higher prices, and houses with more square footage sit on larger lots.

II. An education economist regresses TestScore on ClassSize alone, omitting TeacherQuality. Better teachers raise scores, but whether districts assign their strongest teachers to larger or to smaller classes is not known.

Which statement about the two omitted variable biases is correct?

(a) Both biases are positive

(b) Both biases are negative

(c) The bias in I is positive; the bias in II cannot be signed without knowing the sign of $\text{Cov}(\text{ClassSize}, \text{TeacherQuality})$

(d) The bias in I cannot be signed; the bias in II is negative

(e) Neither estimator is biased, because the omitted variable is absent from the model in both cases

Source: Econ 103 Final, Summer 2026 Session A, Question 9

Question 2, Step 1: Write the Omitted Variable Bias Formula

True model: y depends on two variables:

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

Question 2, Step 1: Write the Omitted Variable Bias Formula

True model: y depends on two variables:

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

Estimated model: the researcher leaves out x_3 and regresses y on x_2 only. Call the estimated slope b_2.

Question 2, Step 1: Write the Omitted Variable Bias Formula

True model: y depends on two variables:

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

Estimated model: the researcher leaves out x_3 and regresses y on x_2 only. Call the estimated slope b_2.
Bias of b_2: the difference between what b_2 estimates on average and the true β_2:

$$\text{bias}(b_2) = E(b_2) - \beta_2 = \beta_3 \times \frac{\text{Cov}(x_2, x_3)}{\text{Var}(x_2)}$$

Question 2, Step 1: Write the Omitted Variable Bias Formula

True model: y depends on two variables:

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

Estimated model: the researcher leaves out x_3 and regresses y on x_2 only. Call the estimated slope b_2.
Bias of b_2: the difference between what b_2 estimates on average and the true β_2:

$$\text{bias}(b_2) = E(b_2) - \beta_2 = \beta_3 \times \frac{\text{Cov}(x_2, x_3)}{\text{Var}(x_2)}$$

The three pieces:

- β_3: the effect of the omitted variable x_3 on y
- $\text{Cov}(x_2, x_3)$: whether the omitted variable x_3 moves with the included variable x_2
- $\text{Var}(x_2)$: the variance of the included variable

Question 2, Step 2: The Rule for the Sign of the Bias

Start from the formula:

$$\text{bias}(b_2) = \beta_3 \times \frac{\text{Cov}(x_2, x_3)}{\text{Var}(x_2)}$$

Question 2, Step 2: The Rule for the Sign of the Bias

Start from the formula:

$$\text{bias}(b_2) = \beta_3 \times \frac{\text{Cov}(x_2, x_3)}{\text{Var}(x_2)}$$

The denominator is always positive: a variance is never negative, and $\text{Var}(x_2) > 0$ whenever x_2 varies. Dividing by a positive number does not change a sign.

Question 2, Step 2: The Rule for the Sign of the Bias

Start from the formula:

$$\text{bias}(b_2) = \beta_3 \times \frac{\text{Cov}(x_2, x_3)}{\text{Var}(x_2)}$$

The denominator is always positive: a variance is never negative, and $\text{Var}(x_2) > 0$ whenever x_2 varies. Dividing by a positive number does not change a sign.
So the sign of the bias is the sign of the product $\beta_3 \times \text{Cov}(x_2, x_3)$:

	$\text{Cov}(x_2, x_3) > 0$	$\text{Cov}(x_2, x_3) < 0$
$\beta_3 > 0$	$(+)(+)$: positive bias	$(+)(-)$: negative bias
$\beta_3 < 0$	$(-)(+)$: negative bias	$(-)(-)$: positive bias

Question 2, Step 2: The Rule for the Sign of the Bias

Start from the formula:

$$\text{bias}(b_2) = \beta_3 \times \frac{\text{Cov}(x_2, x_3)}{\text{Var}(x_2)}$$

The denominator is always positive: a variance is never negative, and $\text{Var}(x_2) > 0$ whenever x_2 varies. Dividing by a positive number does not change a sign.
So the sign of the bias is the sign of the product $\beta_3 \times \text{Cov}(x_2, x_3)$:

	$\text{Cov}(x_2, x_3) > 0$	$\text{Cov}(x_2, x_3) < 0$
$\beta_3 > 0$	$(+)(+)$: positive bias	$(+)(-)$: negative bias
$\beta_3 < 0$	$(-)(+)$: negative bias	$(-)(-)$: positive bias

To sign the bias, we need the sign of both β_3 **and** $\text{Cov}(x_2, x_3)$. If either sign is unknown, the sign of the bias is unknown.

Question 2, Step 3: Case I (Housing)

Label the variables: $y = \text{Price}$, included $x_2 = \text{SQFT}$, omitted $x_3 = \text{LotSize}$.

Question 2, Step 3: Case I (Housing)

Label the variables: $y = \text{Price}$, included $x_2 = \text{SQFT}$, omitted $x_3 = \text{LotSize}$.
Sign of β_3 (effect of LotSize on Price): "Larger lots command higher prices."

$$\beta_3 > 0$$

Question 2, Step 3: Case I (Housing)

Label the variables: $y = \text{Price}$, included $x_2 = \text{SQFT}$, omitted $x_3 = \text{LotSize}$.
Sign of β_3 (effect of LotSize on Price): "Larger lots command higher prices."

$$\beta_3 > 0$$

Sign of $\text{Cov}(\textbf{SQFT}, \textbf{LotSize})$**:** "Houses with more square footage sit on larger lots."

$$\text{Cov}(\text{SQFT}, \text{LotSize}) > 0$$

Question 2, Step 3: Case I (Housing)

Label the variables: $y = \text{Price}$, included $x_2 = \text{SQFT}$, omitted $x_3 = \text{LotSize}$.
Sign of β_3 (effect of LotSize on Price): "Larger lots command higher prices."

$$\beta_3 > 0$$

Sign of $\text{Cov}(\textbf{SQFT}, \textbf{LotSize})$**:** "Houses with more square footage sit on larger lots."

$$\text{Cov}(\text{SQFT}, \text{LotSize}) > 0$$

Sign of the bias:

$$\text{bias}(b_2) = \underbrace{\beta_3}_{+} \times \frac{\overbrace{\text{Cov}(\text{SQFT}, \text{LotSize})}^{+}}{\underbrace{\text{Var}(\text{SQFT})}_{+}} > 0$$

Question 2, Step 3: Case I (Housing)

Label the variables: $y = \text{Price}$, included $x_2 = \text{SQFT}$, omitted $x_3 = \text{LotSize}$.
Sign of β_3 (effect of LotSize on Price): "Larger lots command higher prices."

$$\beta_3 > 0$$

Sign of $\text{Cov}(\textbf{SQFT}, \textbf{LotSize})$**:** "Houses with more square footage sit on larger lots."

$$\text{Cov}(\text{SQFT}, \text{LotSize}) > 0$$

Sign of the bias:

$$\text{bias}(b_2) = \underbrace{\beta_3}_{+} \times \frac{\overbrace{\text{Cov}(\text{SQFT}, \text{LotSize})}^{+}}{\underbrace{\text{Var}(\text{SQFT})}_{+}} > 0$$

Case I: the bias is positive. In words: bigger houses also have bigger lots, so the SQFT coefficient includes part of the price effect of the larger lot. b_2 overstates the effect of square footage.

Question 2, Step 4: Case II (Education)

Label the variables: $y = \text{TestScore}$, included $x_2 = \text{ClassSize}$, omitted $x_3 = \text{TeacherQuality}$.

Question 2, Step 4: Case II (Education)

Label the variables: $y =$ TestScore, included $x_2 =$ ClassSize, omitted $x_3 =$ TeacherQuality.
Sign of β_3 (effect of TeacherQuality on TestScore): "Better teachers raise scores."

$$\beta_3 > 0$$

Question 2, Step 4: Case II (Education)

Label the variables: y = TestScore, included x_2 = ClassSize, omitted x_3 = TeacherQuality.
Sign of β_3 (effect of TeacherQuality on TestScore): "Better teachers raise scores."

$$\beta_3 > 0$$

Sign of Cov(**ClassSize**, **TeacherQuality**)**:** "Whether districts assign their strongest teachers to larger or to smaller classes is not known."

- If the best teachers get the smaller classes: Cov < 0
- If the best teachers get the larger classes: Cov > 0

The question does not tell us which $\implies$ the sign of Cov is unknown.

Question 2, Step 4: Case II (Education)

Label the variables: y = TestScore, included x_2 = ClassSize, omitted x_3 = TeacherQuality.
Sign of β_3 (effect of TeacherQuality on TestScore): "Better teachers raise scores."

$$\beta_3 > 0$$

Sign of Cov(**ClassSize**, **TeacherQuality**)**:** "Whether districts assign their strongest teachers to larger or to smaller classes is not known."

- If the best teachers get the smaller classes: Cov < 0
- If the best teachers get the larger classes: Cov > 0

The question does not tell us which $\implies$ the sign of Cov is unknown.
Sign of the bias:

$$\text{bias}(b_2) = \underbrace{\beta_3}_{+} \times \frac{\overbrace{\text{Cov}(\text{ClassSize}, \text{TeacherQuality})}^{?}}{\underbrace{\text{Var}(\text{ClassSize})}_{+}} = ?$$

Question 2, Step 4: Case II (Education)

Label the variables: y = TestScore, included x_2 = ClassSize, omitted x_3 = TeacherQuality.
Sign of β_3 (effect of TeacherQuality on TestScore): "Better teachers raise scores."

$$\beta_3 > 0$$

Sign of Cov(**ClassSize**, **TeacherQuality**)**:** "Whether districts assign their strongest teachers to larger or to smaller classes is not known."

- If the best teachers get the smaller classes: Cov < 0
- If the best teachers get the larger classes: Cov > 0

The question does not tell us which $\implies$ the sign of Cov is unknown.
Sign of the bias:

$$\text{bias}(b_2) = \underbrace{\beta_3}_{+} \times \frac{\overbrace{\text{Cov}(\text{ClassSize}, \text{TeacherQuality})}^{?}}{\underbrace{\text{Var}(\text{ClassSize})}_{+}} = ?$$

Case II: the bias cannot be signed without knowing the sign of Cov(ClassSize, TeacherQuality).

Question 2, Step 5: Compare with Each Answer Choice

From Steps 3 and 4: Case I is positive. Case II cannot be signed.

Question 2, Step 5: Compare with Each Answer Choice

From Steps 3 and 4: Case I is positive. Case II cannot be signed.

Answer choice	Matches Steps 3 and 4?
(a) Both biases are positive	No: Case II cannot be signed
(b) Both biases are negative	No: Case I is positive
(c) I is positive; II cannot be signed without the sign of Cov(ClassSize, TeacherQuality)	**Yes: both parts match**
(d) I cannot be signed; II is negative	No: both parts are reversed
(e) Neither estimator is biased	No: omitting x_3 causes the bias in the formula of Step 1

Question 2, Step 6: Conclusion

Conclusion

- **Case I (housing):** $\beta_3 > 0$ and $\text{Cov}(\text{SQFT}, \text{LotSize}) > 0$, so the bias is $(+) \times (+)$: **positive**. The regression of Price on SQFT alone overstates the effect of square footage.
- **Case II (education):** $\beta_3 > 0$, but the sign of $\text{Cov}(\text{ClassSize}, \text{TeacherQuality})$ is unknown, so the bias is $(+) \times (?)$: **it cannot be signed**.
- **Answer: (c).**

Outline

1 Question 1: One-Sided t-Test

2 Question 2: Signing Omitted Variable Bias

3 Question 3: Auxiliary Regressions and the VIF

4 Question 4: Testing a Linear Combination

Question 3: An Auxiliary Regression

A researcher is interested in

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + e,$$

but before estimating it, they run

$$X_4 = \alpha_1 + \alpha_2 X_2 + \alpha_3 X_3 + \nu$$

and obtain an R^2 of 0.95.

(a) What is the researcher checking with this second regression?

(b) What should they conclude, and what does it imply for the estimate of β_4?

Source: Econ 103 Final, Summer 2026 Session A, Question 27

Question 3, Step 1: Identify the Second Regression

Look at the left-hand side of each regression:

- Model of interest: the dependent variable is Y.
- Second regression: the dependent variable is X_4, which is one of the *regressors* in the model of interest.

Question 3, Step 1: Identify the Second Regression

Look at the left-hand side of each regression:

- Model of interest: the dependent variable is Y.
- Second regression: the dependent variable is X_4, which is one of the *regressors* in the model of interest.

Look at the right-hand side of the second regression: X_2 and X_3, the *other* regressors in the model of interest.

Question 3, Step 1: Identify the Second Regression

Look at the left-hand side of each regression:

- Model of interest: the dependent variable is Y.
- Second regression: the dependent variable is X_4, which is one of the *regressors* in the model of interest.

Look at the right-hand side of the second regression: X_2 and X_3, the *other* regressors in the model of interest.
Name: a regression of one regressor on the other regressors is an **auxiliary regression**. It is not a model of Y. It describes how the regressors relate to each other.

Question 3, Step 2: Interpret the Auxiliary R^2

What R^2 measures: the share of the variation in the dependent variable that the regressors explain.

Question 3, Step 2: Interpret the Auxiliary R^2

What R^2 measures: the share of the variation in the dependent variable that the regressors explain.
In the auxiliary regression, the dependent variable is X_4 and the regressors are X_2 and X_3. So:

$$R^2 = \text{share of the variation in } X_4 \text{ explained by } X_2 \text{ and } X_3$$

Question 3, Step 2: Interpret the Auxiliary R^2

What R^2 measures: the share of the variation in the dependent variable that the regressors explain.
In the auxiliary regression, the dependent variable is X_4 and the regressors are X_2 and X_3. So:

$$R^2 = \text{share of the variation in } X_4 \text{ explained by } X_2 \text{ and } X_3$$

How to read it:

- R^2 near 0: X_4 has variation of its own, unrelated to X_2 and X_3.
- R^2 near 1: X_4 is almost an exact linear combination of X_2 and X_3.

Question 3, Step 2: Interpret the Auxiliary R^2

What R^2 measures: the share of the variation in the dependent variable that the regressors explain.
In the auxiliary regression, the dependent variable is X_4 and the regressors are X_2 and X_3. So:

$$R^2 = \text{share of the variation in } X_4 \text{ explained by } X_2 \text{ and } X_3$$

How to read it:

- R^2 near 0: X_4 has variation of its own, unrelated to X_2 and X_3.
- R^2 near 1: X_4 is almost an exact linear combination of X_2 and X_3.

Here: $R^2 = 0.95 \implies$ 95% of the variation in X_4 is explained by X_2 and X_3. Only $1 - 0.95 = 0.05$, or 5%, is separate from them.

Question 3, Step 3: Conclusion for Part (a)

Conclusion for (a)

- The second regression is an **auxiliary regression** of the regressor X_4 on the other regressors, X_2 and X_3.
- The researcher is checking for **collinearity**: how much of the variation in X_4 is already explained by X_2 and X_3.

Question 3, Step 4: The Variance of b_4

In a multiple regression, the variance of b_4 is

$$\text{Var}(b_4) = \frac{\sigma^2}{\sum(X_{i4} - \bar{X}_4)^2} \times \frac{1}{1 - R_4^2}$$

where R_4^2 is the R^2 from the auxiliary regression of X_4 on X_2 and X_3.

Question 3, Step 4: The Variance of b_4

In a multiple regression, the variance of b_4 is

$$\text{Var}(b_4) = \frac{\sigma^2}{\sum(X_{i4} - \bar{X}_4)^2} \times \frac{1}{1 - R_4^2}$$

where R_4^2 is the R^2 from the auxiliary regression of X_4 on X_2 and X_3.
The first factor is the variance of b_4 if X_4 were uncorrelated with X_2 and X_3.

Question 3, Step 4: The Variance of b_4

In a multiple regression, the variance of b_4 is

$$\text{Var}(b_4) = \frac{\sigma^2}{\sum(X_{i4} - \bar{X}_4)^2} \times \frac{1}{1 - R_4^2}$$

where R_4^2 is the R^2 from the auxiliary regression of X_4 on X_2 and X_3.
The first factor is the variance of b_4 if X_4 were uncorrelated with X_2 and X_3.
The second factor is the **variance inflation factor** (VIF):

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

It is the number of times larger $\text{Var}(b_4)$ is because of the correlation between X_4 and the other regressors.

Question 3, Step 5: Compute the VIF

Formula:

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

Question 3, Step 5: Compute the VIF

Formula:

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

Plug in $R_4^2 = 0.95$:

$$\text{VIF} = \frac{1}{1 - 0.95}$$

Question 3, Step 5: Compute the VIF

Formula:

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

Plug in $R_4^2 = 0.95$:

$$\text{VIF} = \frac{1}{1 - 0.95}$$

Compute the denominator:

$$1 - 0.95 = 0.05$$

Question 3, Step 5: Compute the VIF

Formula:

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

Plug in $R_4^2 = 0.95$:

$$\text{VIF} = \frac{1}{1 - 0.95}$$

Compute the denominator:

$$1 - 0.95 = 0.05$$

Compute the VIF:

$$\text{VIF} = \frac{1}{0.05} = 20$$

Question 3, Step 5: Compute the VIF

Formula:

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

Plug in $R_4^2 = 0.95$:

$$\text{VIF} = \frac{1}{1 - 0.95}$$

Compute the denominator:

$$1 - 0.95 = 0.05$$

Compute the VIF:

$$\text{VIF} = \frac{1}{0.05} = 20$$

In words: $\text{Var}(b_4)$ is 20 times what it would be if X_4 were uncorrelated with X_2 and X_3.

Question 3, Step 5: Compute the VIF

Formula:

$$\text{VIF} = \frac{1}{1 - R_4^2}$$

Plug in $R_4^2 = 0.95$:

$$\text{VIF} = \frac{1}{1 - 0.95}$$

Compute the denominator:

$$1 - 0.95 = 0.05$$

Compute the VIF:

$$\text{VIF} = \frac{1}{0.05} = 20$$

In words: $\text{Var}(b_4)$ is 20 times what it would be if X_4 were uncorrelated with X_2 and X_3.
Rule of thumb: a VIF above 5 or above 10 indicates a collinearity problem. A VIF of 20 is above both.

Question 3, Step 6: Effect on the Standard Error of b_4

The standard error is the square root of the variance. If the variance is 20 times larger, the standard error is $\sqrt{20}$ times larger:

$$\text{se}(b_4) = \sqrt{\text{Var}(b_4)} \qquad \Longrightarrow \qquad \text{factor} = \sqrt{20}$$

Question 3, Step 6: Effect on the Standard Error of b_4

The standard error is the square root of the variance. If the variance is 20 times larger, the standard error is $\sqrt{20}$ times larger:

$$\text{se}(b_4) = \sqrt{\text{Var}(b_4)} \qquad \Longrightarrow \qquad \text{factor} = \sqrt{20}$$

Compute:

$$\sqrt{20} = 4.47$$

Question 3, Step 6: Effect on the Standard Error of b_4

The standard error is the square root of the variance. If the variance is 20 times larger, the standard error is $\sqrt{20}$ times larger:

$$\text{se}(b_4) = \sqrt{\text{Var}(b_4)} \qquad \Longrightarrow \qquad \text{factor} = \sqrt{20}$$

Compute:

$$\sqrt{20} = 4.47$$

In words: the standard error of b_4 is about 4.5 times larger than it would be without the collinearity.

Question 3, Step 6: Effect on the Standard Error of b_4

The standard error is the square root of the variance. If the variance is 20 times larger, the standard error is $\sqrt{20}$ times larger:

$$\text{se}(b_4) = \sqrt{\text{Var}(b_4)} \qquad \Longrightarrow \qquad \text{factor} = \sqrt{20}$$

Compute:

$$\sqrt{20} = 4.47$$

In words: the standard error of b_4 is about 4.5 times larger than it would be without the collinearity.
Consequences of a larger standard error:

- The t-statistic $t = b_4/\text{se}(b_4)$ is smaller $\Longrightarrow$ it is harder to reject $H_0 \colon \beta_4 = 0$.
- The confidence interval $b_4 \pm t_c \times \text{se}(b_4)$ is wider.

Question 3, Step 7: Is b_4 Biased?

Unbiasedness requires $E(e \mid X_2, X_3, X_4) = 0$. This assumption is about the error term. It says nothing about how the regressors are correlated with each other.

Question 3, Step 7: Is b_4 Biased?

Unbiasedness requires $E(e \mid X_2, X_3, X_4) = 0$. This assumption is about the error term. It says nothing about how the regressors are correlated with each other.

Collinearity is a correlation among the regressors. It does not change the error term. So

$$E(b_4) = \beta_4 \qquad \text{still holds.}$$

Question 3, Step 7: Is b_4 Biased?

Unbiasedness requires $E(e \mid X_2, X_3, X_4) = 0$. This assumption is about the error term. It says nothing about how the regressors are correlated with each other.

Collinearity is a correlation among the regressors. It does not change the error term. So

$$E(b_4) = \beta_4 \qquad \text{still holds.}$$

Property of b_4	Affected?	How
Unbiasedness	No	$E(b_4) = \beta_4$
Variance	Yes	20 times larger
Standard error	Yes	4.47 times larger
t-statistic	Yes	smaller
Confidence interval	Yes	wider

Question 3, Step 8: Conclusion for Part (b)

Conclusion for (b)

- **Conclusion:** there is severe collinearity. X_2 and X_3 explain 95% of the variation in X_4.
- **VIF:** $\frac{1}{1-0.95} = 20$, above the rule-of-thumb cutoffs of 5 and 10.
- **Implication for b_4:** b_4 is still unbiased, but its variance is 20 times larger and its standard error is $\sqrt{20} = 4.47$ times larger. β_4 is estimated imprecisely: wide confidence intervals and small t-statistics.

Outline

1. Question 1: One-Sided t-Test
2. Question 2: Signing Omitted Variable Bias
3. Question 3: Auxiliary Regressions and the VIF
4. Question 4: Testing a Linear Combination

Question 4: Studying and Coffee

From a random sample of 52 students we collect test scores, hours studied, and cups of coffee consumed, and estimate

$$\widehat{\text{Score}} = 50 + 6\,\text{Study} + 3\,\text{Coffee} + 0.5\,(\text{Study} \times \text{Coffee})$$

The estimated variances (diagonal) and covariances (off-diagonal) of (b_1, b_2, b_3, b_4) are

	b_1	b_2	b_3	b_4
b_1	4	0.2	-1.5	-0.2
b_2	0.2	4	0.3	-0.75
b_3	-1.5	0.3	2	-0.1
b_4	-0.2	-0.75	-0.1	0.75

Consider the marginal effect of another hour of studying for a student who has drunk 2 cups of coffee.

(a) Test, at $\alpha = 0.05$, whether that marginal effect is more than 4 points, against the null that it is at most 4.

(b) Construct the 90% confidence interval for that marginal effect.

Source: Econ 103 Final, Summer 2026 Session A, Question 24

Question 4, Step 1: Find the Marginal Effect of Study

The model:

$$E(\text{Score}) = \beta_1 + \beta_2\,\text{Study} + \beta_3\,\text{Coffee} + \beta_4\,(\text{Study} \times \text{Coffee})$$

Question 4, Step 1: Find the Marginal Effect of Study

The model:

$$E(\text{Score}) = \beta_1 + \beta_2\,\text{Study} + \beta_3\,\text{Coffee} + \beta_4\,(\text{Study} \times \text{Coffee})$$

The marginal effect of Study is the derivative with respect to Study. Take it term by term:

Term	Derivative with respect to Study	Reason
β_1	0	constant
$\beta_2\,\text{Study}$	β_2	
$\beta_3\,\text{Coffee}$	0	does not contain Study
$\beta_4\,\text{Study} \times \text{Coffee}$	$\beta_4\,\text{Coffee}$	Coffee is held fixed

Question 4, Step 1: Find the Marginal Effect of Study

The model:

$$E(\text{Score}) = \beta_1 + \beta_2\,\text{Study} + \beta_3\,\text{Coffee} + \beta_4\,(\text{Study} \times \text{Coffee})$$

The marginal effect of Study is the derivative with respect to Study. Take it term by term:

Term	Derivative with respect to Study	Reason
β_1	0	constant
$\beta_2\,\text{Study}$	β_2	
$\beta_3\,\text{Coffee}$	0	does not contain Study
$\beta_4\,\text{Study} \times \text{Coffee}$	$\beta_4\,\text{Coffee}$	Coffee is held fixed

Add the terms:

$$\frac{\partial\, E(\text{Score})}{\partial\,\text{Study}} = \beta_2 + \beta_4\,\text{Coffee}$$

Question 4, Step 1: Find the Marginal Effect of Study

The model:

$$E(\text{Score}) = \beta_1 + \beta_2\,\text{Study} + \beta_3\,\text{Coffee} + \beta_4\,(\text{Study} \times \text{Coffee})$$

The marginal effect of Study is the derivative with respect to Study. Take it term by term:

Term	Derivative with respect to Study	Reason
β_1	0	constant
$\beta_2\,\text{Study}$	β_2	
$\beta_3\,\text{Coffee}$	0	does not contain Study
$\beta_4\,\text{Study} \times \text{Coffee}$	$\beta_4\,\text{Coffee}$	Coffee is held fixed

Add the terms:

$$\frac{\partial\, E(\text{Score})}{\partial\,\text{Study}} = \beta_2 + \beta_4\,\text{Coffee}$$

In words: the effect of one more hour of studying depends on how much coffee the student drank.

Question 4, Step 2: Evaluate at Coffee = 2

Marginal effect from Step 1:

$$\beta_2 + \beta_4 \, \text{Coffee}$$

Question 4, Step 2: Evaluate at Coffee $= 2$

Marginal effect from Step 1:

$$\beta_2 + \beta_4 \, \text{Coffee}$$

Substitute Coffee $= 2$:

$$\beta_2 + \beta_4 \times 2 = \beta_2 + 2\beta_4$$

Question 4, Step 2: Evaluate at Coffee $= 2$

Marginal effect from Step 1:

$$\beta_2 + \beta_4 \, \text{Coffee}$$

Substitute Coffee $= 2$:

$$\beta_2 + \beta_4 \times 2 = \beta_2 + 2\beta_4$$

Give it a name: the quantity we want to test and estimate is

$$\theta = \beta_2 + 2\beta_4$$

θ is a **linear combination** of two coefficients: 1 times β_2 plus 2 times β_4.

Question 4, Step 3: Compute the Point Estimate

Replace each β with its estimate:

$$\hat{\theta} = b_2 + 2b_4$$

Question 4, Step 3: Compute the Point Estimate

Replace each β with its estimate:

$$\hat{\theta} = b_2 + 2b_4$$

Read the estimates from the fitted regression:

$$\widehat{\text{Score}} = 50 + \underbrace{6}_{b_2}\,\text{Study} + 3\,\text{Coffee} + \underbrace{0.5}_{b_4}\,(\text{Study} \times \text{Coffee})$$

Question 4, Step 3: Compute the Point Estimate

Replace each β with its estimate:

$$\hat{\theta} = b_2 + 2b_4$$

Read the estimates from the fitted regression:

$$\widehat{\text{Score}} = 50 + \underbrace{6}_{b_2}\,\text{Study} + 3\,\text{Coffee} + \underbrace{0.5}_{b_4}\,(\text{Study} \times \text{Coffee})$$

Plug in:

$$\hat{\theta} = 6 + 2 \times 0.5$$

Question 4, Step 3: Compute the Point Estimate

Replace each β with its estimate:

$$\hat{\theta} = b_2 + 2b_4$$

Read the estimates from the fitted regression:

$$\widehat{\text{Score}} = 50 + \underbrace{6}_{b_2}\,\text{Study} + 3\,\text{Coffee} + \underbrace{0.5}_{b_4}\,(\text{Study} \times \text{Coffee})$$

Plug in:

$$\hat{\theta} = 6 + 2 \times 0.5$$

Compute:

$$2 \times 0.5 = 1 \qquad \Longrightarrow \qquad \hat{\theta} = 6 + 1 = 7$$

Question 4, Step 3: Compute the Point Estimate

Replace each β with its estimate:

$$\hat{\theta} = b_2 + 2b_4$$

Read the estimates from the fitted regression:

$$\widehat{\text{Score}} = 50 + \underbrace{6}_{b_2}\,\text{Study} + 3\,\text{Coffee} + \underbrace{0.5}_{b_4}\,(\text{Study} \times \text{Coffee})$$

Plug in:

$$\hat{\theta} = 6 + 2 \times 0.5$$

Compute:

$$2 \times 0.5 = 1 \qquad \Longrightarrow \qquad \hat{\theta} = 6 + 1 = 7$$

In words: for a student who drank 2 cups of coffee, one more hour of studying is associated with a test score 7 points higher.

Question 4, Step 4: Read the Entries We Need from the Table

$\hat{\theta} = b_2 + 2b_4$ **involves** b_2 **and** b_4**.** Its variance needs $\text{Var}(b_2)$, $\text{Var}(b_4)$, and $\text{Cov}(b_2, b_4)$.

	b_1	b_2	b_3	b_4
b_1	4	0.2	-1.5	-0.2
b_2	0.2	**4**	0.3	$\mathbf{-0.75}$
b_3	-1.5	0.3	2	-0.1
b_4	-0.2	-0.75	-0.1	**0.75**

Question 4, Step 4: Read the Entries We Need from the Table

$\hat{\theta} = b_2 + 2b_4$ **involves** b_2 **and** b_4**.** Its variance needs $\text{Var}(b_2)$, $\text{Var}(b_4)$, and $\text{Cov}(b_2, b_4)$.

	b_1	b_2	b_3	b_4
b_1	4	0.2	-1.5	-0.2
b_2	0.2	**4**	0.3	$\mathbf{-0.75}$
b_3	-1.5	0.3	2	-0.1
b_4	-0.2	-0.75	-0.1	**0.75**

- $\text{Var}(b_2)$: row b_2, column b_2 (diagonal) $= 4$
- $\text{Var}(b_4)$: row b_4, column b_4 (diagonal) $= 0.75$
- $\text{Cov}(b_2, b_4)$: row b_2, column b_4 (off-diagonal) $= -0.75$

Question 4, Step 5: Compute the Variance of $\hat{\theta}$

Variance of a linear combination $a\,b_2 + c\,b_4$:

$$\text{Var}(a\,b_2 + c\,b_4) = a^2\,\text{Var}(b_2) + c^2\,\text{Var}(b_4) + 2\,a\,c\;\text{Cov}(b_2, b_4)$$

Here $a = 1$ (the multiplier on b_2) and $c = 2$ (the multiplier on b_4).

Question 4, Step 5: Compute the Variance of $\hat{\theta}$

Variance of a linear combination $a\,b_2 + c\,b_4$:

$$\text{Var}(a\,b_2 + c\,b_4) = a^2\,\text{Var}(b_2) + c^2\,\text{Var}(b_4) + 2\,a\,c\;\text{Cov}(b_2, b_4)$$

Here $a = 1$ (the multiplier on b_2) and $c = 2$ (the multiplier on b_4).
Term 1: $a^2\,\text{Var}(b_2) = 1^2 \times 4 = 1 \times 4 = 4$

Question 4, Step 5: Compute the Variance of $\hat{\theta}$

Variance of a linear combination $a\,b_2 + c\,b_4$:

$$\text{Var}(a\,b_2 + c\,b_4) = a^2\,\text{Var}(b_2) + c^2\,\text{Var}(b_4) + 2\,a\,c\;\text{Cov}(b_2, b_4)$$

Here $a = 1$ (the multiplier on b_2) and $c = 2$ (the multiplier on b_4).
Term 1: $a^2\,\text{Var}(b_2) = 1^2 \times 4 = 1 \times 4 = 4$
Term 2: $c^2\,\text{Var}(b_4) = 2^2 \times 0.75 = 4 \times 0.75 = 3$

Question 4, Step 5: Compute the Variance of $\hat{\theta}$

Variance of a linear combination $a\,b_2 + c\,b_4$:

$$\text{Var}(a\,b_2 + c\,b_4) = a^2\,\text{Var}(b_2) + c^2\,\text{Var}(b_4) + 2\,a\,c\;\text{Cov}(b_2, b_4)$$

Here $a = 1$ (the multiplier on b_2) and $c = 2$ (the multiplier on b_4).
Term 1: $a^2\,\text{Var}(b_2) = 1^2 \times 4 = 1 \times 4 = 4$
Term 2: $c^2\,\text{Var}(b_4) = 2^2 \times 0.75 = 4 \times 0.75 = 3$
Term 3: $2\,a\,c\;\text{Cov}(b_2, b_4) = 2 \times 1 \times 2 \times (-0.75) = 4 \times (-0.75) = -3$

Question 4, Step 5: Compute the Variance of $\hat{\theta}$

Variance of a linear combination $a\,b_2 + c\,b_4$:

$$\text{Var}(a\,b_2 + c\,b_4) = a^2\,\text{Var}(b_2) + c^2\,\text{Var}(b_4) + 2\,a\,c\;\text{Cov}(b_2, b_4)$$

Here $a = 1$ (the multiplier on b_2) and $c = 2$ (the multiplier on b_4).
Term 1: $a^2\,\text{Var}(b_2) = 1^2 \times 4 = 1 \times 4 = 4$
Term 2: $c^2\,\text{Var}(b_4) = 2^2 \times 0.75 = 4 \times 0.75 = 3$
Term 3: $2\,a\,c\;\text{Cov}(b_2, b_4) = 2 \times 1 \times 2 \times (-0.75) = 4 \times (-0.75) = -3$
Add the three terms:

$$\text{Var}(\hat{\theta}) = 4 + 3 + (-3) = 4$$

Question 4, Step 6: Compute the Standard Error of $\hat{\theta}$

The standard error is the square root of the variance:

$$\text{se}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}$$

Question 4, Step 6: Compute the Standard Error of $\hat{\theta}$

The standard error is the square root of the variance:

$$\text{se}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}$$

Plug in $\text{Var}(\hat{\theta}) = 4$ from Step 5:

$$\text{se}(\hat{\theta}) = \sqrt{4} = 2$$

Question 4, Step 6: Compute the Standard Error of $\hat{\theta}$

The standard error is the square root of the variance:

$$\text{se}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}$$

Plug in $\text{Var}(\hat{\theta}) = 4$ from Step 5:

$$\text{se}(\hat{\theta}) = \sqrt{4} = 2$$

Summary so far:

Quantity	Value
Marginal effect at Coffee $= 2$	$\theta = \beta_2 + 2\beta_4$
Point estimate	$\hat{\theta} = 7$
Standard error	$\text{se}(\hat{\theta}) = 2$

Parts (a) and (b) both use these two numbers.

Question 4, Part (a), Step 7: Write the Hypotheses

The question: is the marginal effect *more than* 4 points, against the null that it is *at most* 4?

Question 4, Part (a), Step 7: Write the Hypotheses

The question: is the marginal effect *more than* 4 points, against the null that it is *at most* 4?
Null (at most 4):

$$H_0 : \beta_2 + 2\beta_4 \leq 4$$

Question 4, Part (a), Step 7: Write the Hypotheses

The question: is the marginal effect *more than* 4 points, against the null that it is *at most* 4?
Null (at most 4):

$$H_0 : \beta_2 + 2\beta_4 \leq 4$$

Alternative (more than 4):

$$H_1 : \beta_2 + 2\beta_4 > 4$$

Question 4, Part (a), Step 7: Write the Hypotheses

The question: is the marginal effect *more than* 4 points, against the null that it is *at most* 4?
Null (at most 4):

$$H_0 : \beta_2 + 2\beta_4 \leq 4$$

Alternative (more than 4):

$$H_1 : \beta_2 + 2\beta_4 > 4$$

Which tail? H_1 points to values greater than 4 $\implies$ a **right-tail test**.

Question 4, Part (a), Step 8: Compute the t-Statistic

Formula: compare the estimate to the boundary of H_0, which is 4:

$$t = \frac{\hat{\theta} - 4}{\text{se}(\hat{\theta})}$$

Question 4, Part (a), Step 8: Compute the t-Statistic

Formula: compare the estimate to the boundary of H_0, which is 4:

$$t = \frac{\hat{\theta} - 4}{\text{se}(\hat{\theta})}$$

Plug in $\hat{\theta} = 7$ (Step 3) and $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$t = \frac{7 - 4}{2}$$

Question 4, Part (a), Step 8: Compute the t-Statistic

Formula: compare the estimate to the boundary of H_0, which is 4:

$$t = \frac{\hat{\theta} - 4}{\text{se}(\hat{\theta})}$$

Plug in $\hat{\theta} = 7$ (Step 3) and $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$t = \frac{7 - 4}{2}$$

Compute the numerator: $7 - 4 = 3$

Question 4, Part (a), Step 8: Compute the t-Statistic

Formula: compare the estimate to the boundary of H_0, which is 4:

$$t = \frac{\hat{\theta} - 4}{\text{se}(\hat{\theta})}$$

Plug in $\hat{\theta} = 7$ (Step 3) and $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$t = \frac{7 - 4}{2}$$

Compute the numerator: $7 - 4 = 3$
Compute t:

$$t = \frac{3}{2} = 1.5$$

Question 4, Part (a), Step 8: Compute the t-Statistic

Formula: compare the estimate to the boundary of H_0, which is 4:

$$t = \frac{\hat{\theta} - 4}{\text{se}(\hat{\theta})}$$

Plug in $\hat{\theta} = 7$ (Step 3) and $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$t = \frac{7 - 4}{2}$$

Compute the numerator: $7 - 4 = 3$
Compute t:

$$t = \frac{3}{2} = 1.5$$

In words: the estimate 7 is 1.5 standard errors above 4.

Question 4, Part (a), Step 9: Find the Critical Value

Degrees of freedom: $N - K$, where K counts every coefficient, including the intercept:

$$N = 52, \qquad K = 4 \ (\beta_1, \beta_2, \beta_3, \beta_4) \qquad \Longrightarrow \qquad N - K = 52 - 4 = 48$$

Question 4, Part (a), Step 9: Find the Critical Value

Degrees of freedom: $N - K$, where K counts every coefficient, including the intercept:

$$N = 52, \qquad K = 4 \ (\beta_1, \beta_2, \beta_3, \beta_4) \qquad \Longrightarrow \qquad N - K = 52 - 4 = 48$$

Significance level and tails: $\alpha = 0.05$, one tail (right) $\Longrightarrow$ we need the value with 95% of the t_{48} distribution below it:

$$t_c = t_{(0.95,\ 48)}$$

Question 4, Part (a), Step 9: Find the Critical Value

Degrees of freedom: $N - K$, where K counts every coefficient, including the intercept:

$$N = 52, \qquad K = 4\ (\beta_1, \beta_2, \beta_3, \beta_4) \qquad \Longrightarrow \qquad N - K = 52 - 4 = 48$$

Significance level and tails: $\alpha = 0.05$, one tail (right) $\Longrightarrow$ we need the value with 95% of the t_{48} distribution below it:

$$t_c = t_{(0.95,\ 48)}$$

Read the table: the table prints rows 40 and 50, not 48:

$$t_{(0.95,\ 40)} = 1.684 \qquad t_{(0.95,\ 50)} = 1.676$$

Both round to 1.68, so $t_c = 1.68$.

Question 4, Part (a), Step 9: Find the Critical Value

Degrees of freedom: $N - K$, where K counts every coefficient, including the intercept:

$$N = 52, \qquad K = 4\ (\beta_1, \beta_2, \beta_3, \beta_4) \qquad \Longrightarrow \qquad N - K = 52 - 4 = 48$$

Significance level and tails: $\alpha = 0.05$, one tail (right) $\Longrightarrow$ we need the value with 95% of the t_{48} distribution below it:

$$t_c = t_{(0.95,\ 48)}$$

Read the table: the table prints rows 40 and 50, not 48:

$$t_{(0.95,\ 40)} = 1.684 \qquad t_{(0.95,\ 50)} = 1.676$$

Both round to 1.68, so $t_c = 1.68$.
Rejection rule: reject H_0 if $t > 1.68$.

Question 4, Part (a), Step 10: Make the Decision

Compare the test statistic (Step 8) to the critical value (Step 9):

$$t = 1.5 \qquad \text{vs} \qquad t_c = 1.68$$

Question 4, Part (a), Step 10: Make the Decision

Compare the test statistic (Step 8) to the critical value (Step 9):

$$t = 1.5 \qquad \text{vs} \qquad t_c = 1.68$$

$$1.5 < 1.68$$

Question 4, Part (a), Step 10: Make the Decision

Compare the test statistic (Step 8) to the critical value (Step 9):

$$t = 1.5 \qquad \text{vs} \qquad t_c = 1.68$$

$$1.5 < 1.68$$

The test statistic is not in the rejection region $\implies$ fail to reject $H_0 : \beta_2 + 2\beta_4 \leq 4$.

Question 4, Part (a), Step 11: Conclusion for Part (a)

Conclusion for (a)

- **Decision:** Fail to reject $H_0: \beta_2 + 2\beta_4 \leq 4$ at the 5% level.
- **Evidence:** $t = 1.5 < 1.68$.
- **Answer to the question:** for a student who drank 2 cups of coffee, the estimated effect of one more hour of studying is 7 points, but at the 5% level the data do not provide enough evidence that the effect is more than 4 points.

Question 4, Part (b), Step 12: Find the Critical Value for a 90% Interval

Confidence level: $90\% = 1 - \alpha \implies \alpha = 0.10$.

Question 4, Part (b), Step 12: Find the Critical Value for a 90% Interval

Confidence level: $90\% = 1 - \alpha \implies \alpha = 0.10$.
Split α between the two tails: a confidence interval has two ends, so half of α goes in each tail:

$$\frac{\alpha}{2} = \frac{0.10}{2} = 0.05 \text{ in each tail}$$

Question 4, Part (b), Step 12: Find the Critical Value for a 90% Interval

Confidence level: $90\% = 1 - \alpha \implies \alpha = 0.10$.
Split α between the two tails: a confidence interval has two ends, so half of α goes in each tail:

$$\frac{\alpha}{2} = \frac{0.10}{2} = 0.05 \text{ in each tail}$$

So we need the value with 95% of the distribution below it (5% in the upper tail), with 48 degrees of freedom (Step 9):

$$t_c = t_{(0.95,\ 48)} = 1.68$$

This is the same critical value as in Part (a).

Question 4, Part (b), Step 13: Compute the Interval

Formula:

$$\hat{\theta} \pm t_c \times \text{se}(\hat{\theta})$$

Question 4, Part (b), Step 13: Compute the Interval

Formula:

$$\hat{\theta} \pm t_c \times \text{se}(\hat{\theta})$$

Plug in $\hat{\theta} = 7$ (Step 3), $t_c = 1.68$ (Step 12), $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$7 \pm 1.68 \times 2$$

Question 4, Part (b), Step 13: Compute the Interval

Formula:

$$\hat{\theta} \pm t_c \times \text{se}(\hat{\theta})$$

Plug in $\hat{\theta} = 7$ (Step 3), $t_c = 1.68$ (Step 12), $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$7 \pm 1.68 \times 2$$

Margin of error:

$$1.68 \times 2 = 3.36$$

Question 4, Part (b), Step 13: Compute the Interval

Formula:

$$\hat{\theta} \pm t_c \times \text{se}(\hat{\theta})$$

Plug in $\hat{\theta} = 7$ (Step 3), $t_c = 1.68$ (Step 12), $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$7 \pm 1.68 \times 2$$

Margin of error:

$$1.68 \times 2 = 3.36$$

Lower endpoint: $7 - 3.36 = 3.64$ **Upper endpoint:** $7 + 3.36 = 10.36$

$$[3.64,\ 10.36]$$

Question 4, Part (b), Step 13: Compute the Interval

Formula:

$$\hat{\theta} \pm t_c \times \text{se}(\hat{\theta})$$

Plug in $\hat{\theta} = 7$ (Step 3), $t_c = 1.68$ (Step 12), $\text{se}(\hat{\theta}) = 2$ (Step 6):

$$7 \pm 1.68 \times 2$$

Margin of error:

$$1.68 \times 2 = 3.36$$

Lower endpoint: $7 - 3.36 = 3.64$ **Upper endpoint:** $7 + 3.36 = 10.36$

$$[3.64,\ 10.36]$$

Check against Part (a): the interval contains 4. So the data cannot rule out a marginal effect of 4, which agrees with failing to reject H_0 in Part (a).

Question 4, Part (b), Step 14: Conclusion for Part (b)

Conclusion for (b)

- **90% confidence interval:** $[3.64,\ 10.36]$
- **In words:** for a student who drank 2 cups of coffee, we are 90% confident that one more hour of studying is associated with a test score between 3.64 and 10.36 points higher.

Thank you!
jakeanderson@g.ucla.edu