Linear Regression with One Regressor: Hypothesis Tests and Confidence Intervals

SW Chapter 5

ECON3500: Econometrics and Applications

Spring 2026

Learning Objectives

By the end of this chapter, you will be able to:

Create hypotheses about slope coefficients and test them using $\hat{\beta}_1$ and its standard error

Correctly interpret the results of hypothesis tests

Calculate confidence intervals for β_1

Take binary regressors in stride (and interpret them correctly)

Understand the implications of heteroskedasticity and correct your standard errors

Know and apply the Gauss–Markov theorem to understand the circumstances under which OLS is BLUE

Where We Are Going

We want to learn about the slope of the population regression line. We have data from a sample, so there is sampling uncertainty.

State the population object of interest

Provide an estimator of this population object

Derive the sampling distribution of the estimator (large-n normal by the CLT)

Find the standard error (SE) of the estimator

Construct t-statistics and confidence intervals

5.1 Testing Hypotheses About One Regression Coefficient

The challenge: Sampling uncertainty

Recall that $\hat{\beta}_1$ is a random variable with its own sampling distribution.

We estimated the regression coefficient ($\hat{\beta}_1$) of education on wages.

But this is just **one sample** from the population

If we drew a different sample, we'd get a different $\hat{\beta}_1$

Question: How confident can we be that the true $\beta_1 \neq 0$)?

Review: The Sampling Distribution of $\hat{\beta}_1$

Recall: under the Least Squares Assumptions, for n large, $\hat{\beta}_1$ is approximately distributed:

> **Note**
>
> $$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma_v^2}{n(\sigma_X^2)^2}\right), \quad \text{where } v_i = (X_i - \mu_X)u_i$$

Note: We won't compute variances by hand, but the intuition is useful.

Key insight:

Under 3 LS assumptions, $\hat{\beta}_1$ is centered at the true β_1 (unbiased)

The spread (variance) depends on sample size n, variation in X, and error variance (σ_v^2)

CLT $\Rightarrow$ $\hat{\beta}_1$ is approximately normal in large samples

Hypothesis Testing: General Setup

Null hypothesis and two-sided alternative:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{vs} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

Null hypothesis and one-sided alternative:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{vs} \quad H_1 : \beta_1 < \beta_{1,0}$$

where $\beta_{1,0}$ is the hypothesized value of β_1 under the null (usually 0).

General Approach

General formula for any hypothesis test:

$$t = \frac{\text{estimator} - \text{hypothesized value}}{SE(\text{estimator})}$$

For testing μ: $t = \frac{\bar{Y} - \mu_{Y,0}}{s_y / \sqrt{n}}$

For testing β_1:

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)}$$

Testing $H_0 : \beta_{1,0} = 0$

Construct your t-statistic:

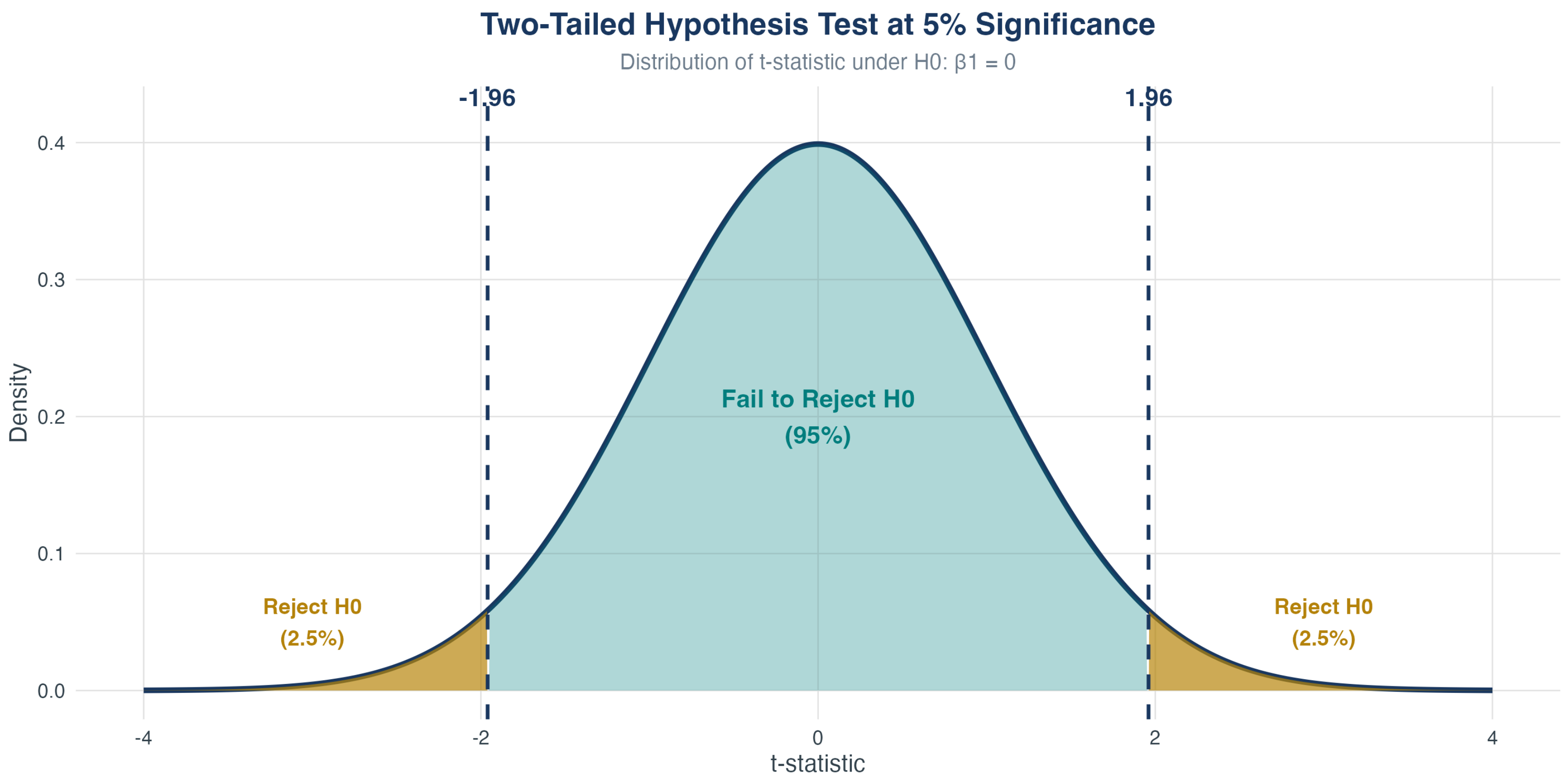
$$t = \frac{\hat{\beta_1} - \beta_{1,0}}{SE(\hat{\beta_1})}$$

Measures how many standard errors $\hat{\beta_1}$ is away from $\beta_{1,0}$

If $|t|$ is large, $\hat{\beta_1}$ is far from $\beta_{1,0}$ $\Rightarrow$ evidence against H_0

Under H_0, $t \sim N(0, 1)$ in large samples (rule of thumb: $n > 30$)

Decision Rule: When to Reject H_0



Reject H_0 at significance level α if:

Choosing a Significance Level

We fix α as the **significance level** (Type I error)

α is the probability of falsely rejecting the null

Typical choice: $\alpha = 0.05$ (5% false positives)

Is that too high? Why not make α super small?

Choosing a Significance Level (Tradeoff)

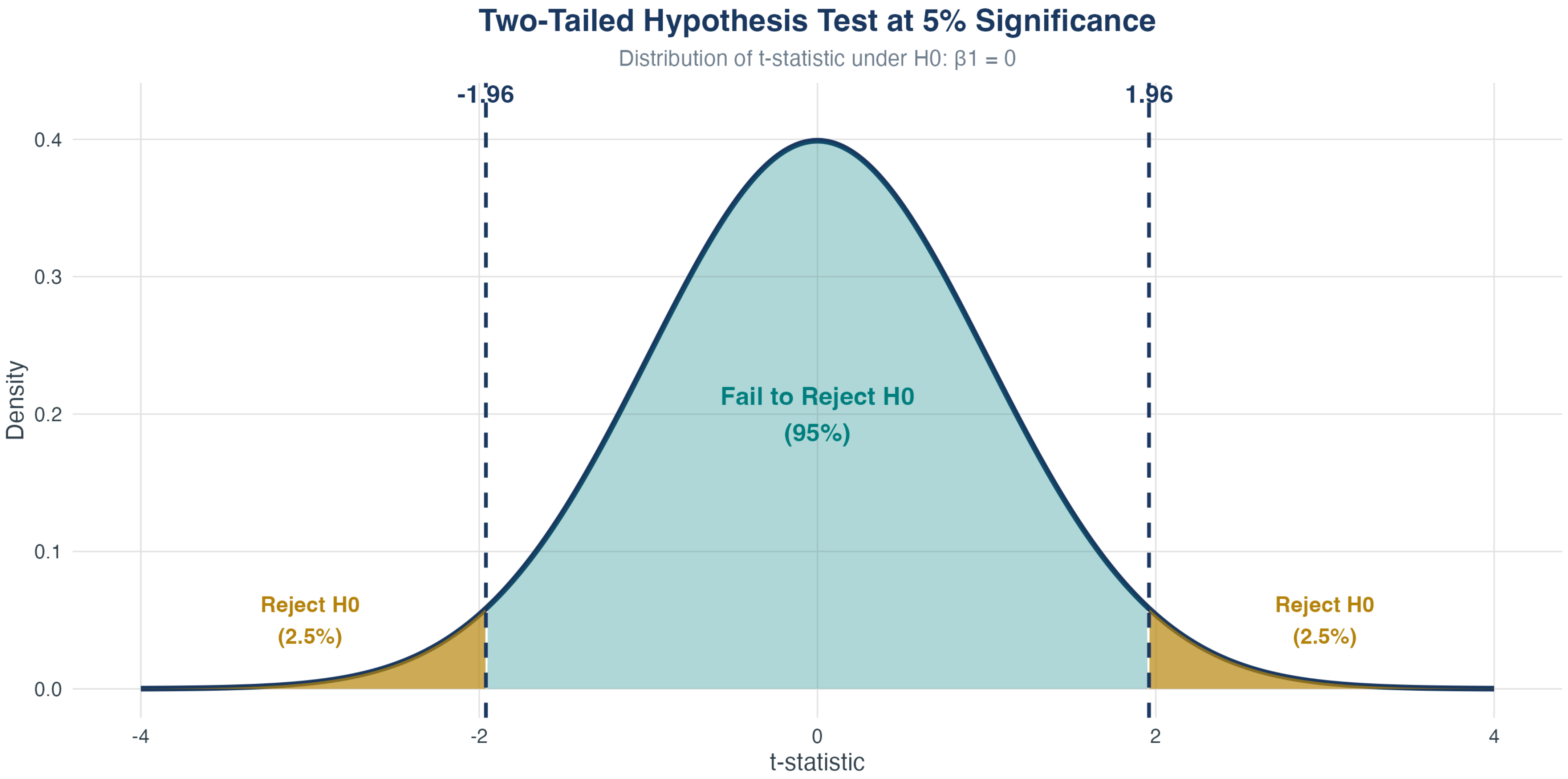
Smaller α makes it harder to reject H_0

Fewer false positives, but more false negatives

Power = $1 - \beta$ (probability of rejecting a false null)

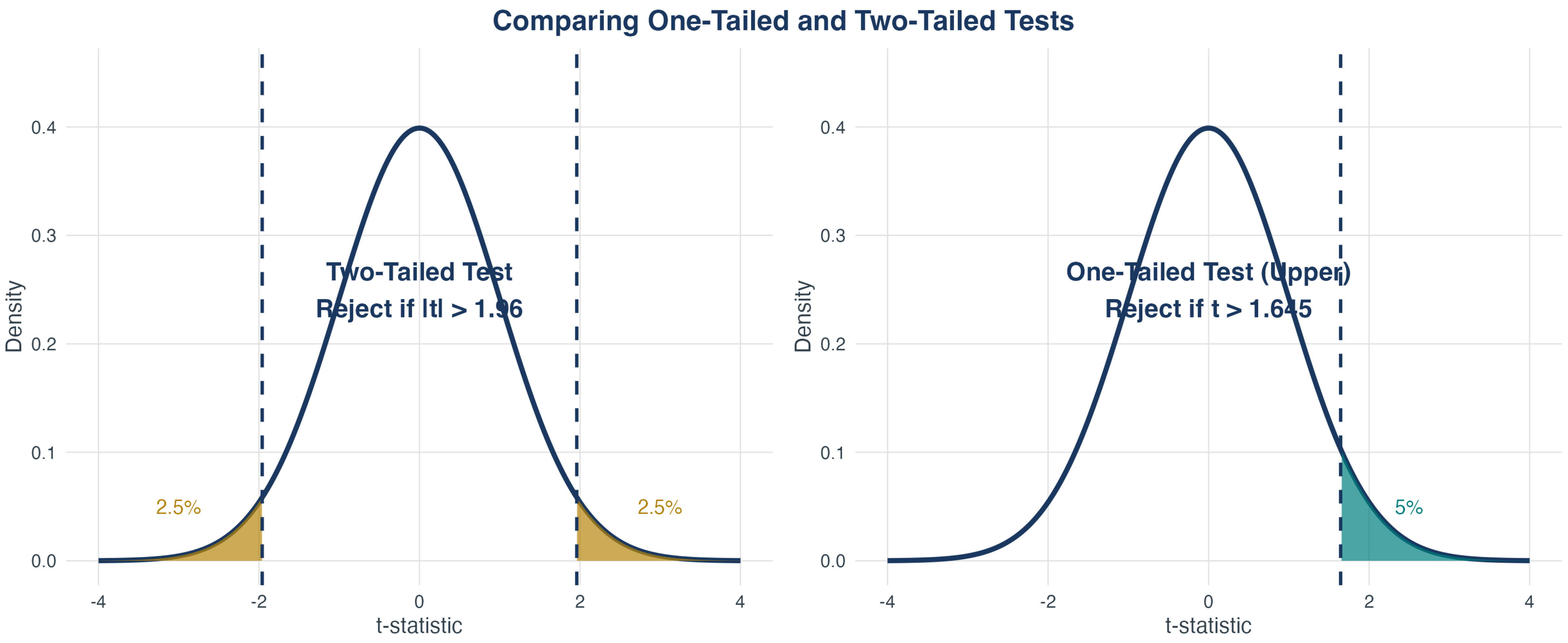
Tradeoff between significance (α) and power (β)

Decision Rule: When to Reject H_0



Rejection regions for one- and two-sided tests

One-Tailed vs Two-Tailed Tests



Which to use?

Critical Values for Common Significance Levels

Significance level	Two-tailed	One-tailed
10%	1.645	1.28
5%	1.96	1.645
1%	2.58	2.33

Rule of thumb: For two-tailed tests at 5%, reject if $|t| > 2$.

Knowledge Check: Calculate the t-Statistic

Question

You estimate $\hat{\beta_1} = 3.5$ with $\mathrm{SE}(\hat{\beta_1}) = 1.2$. Test $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$ at 5% level.

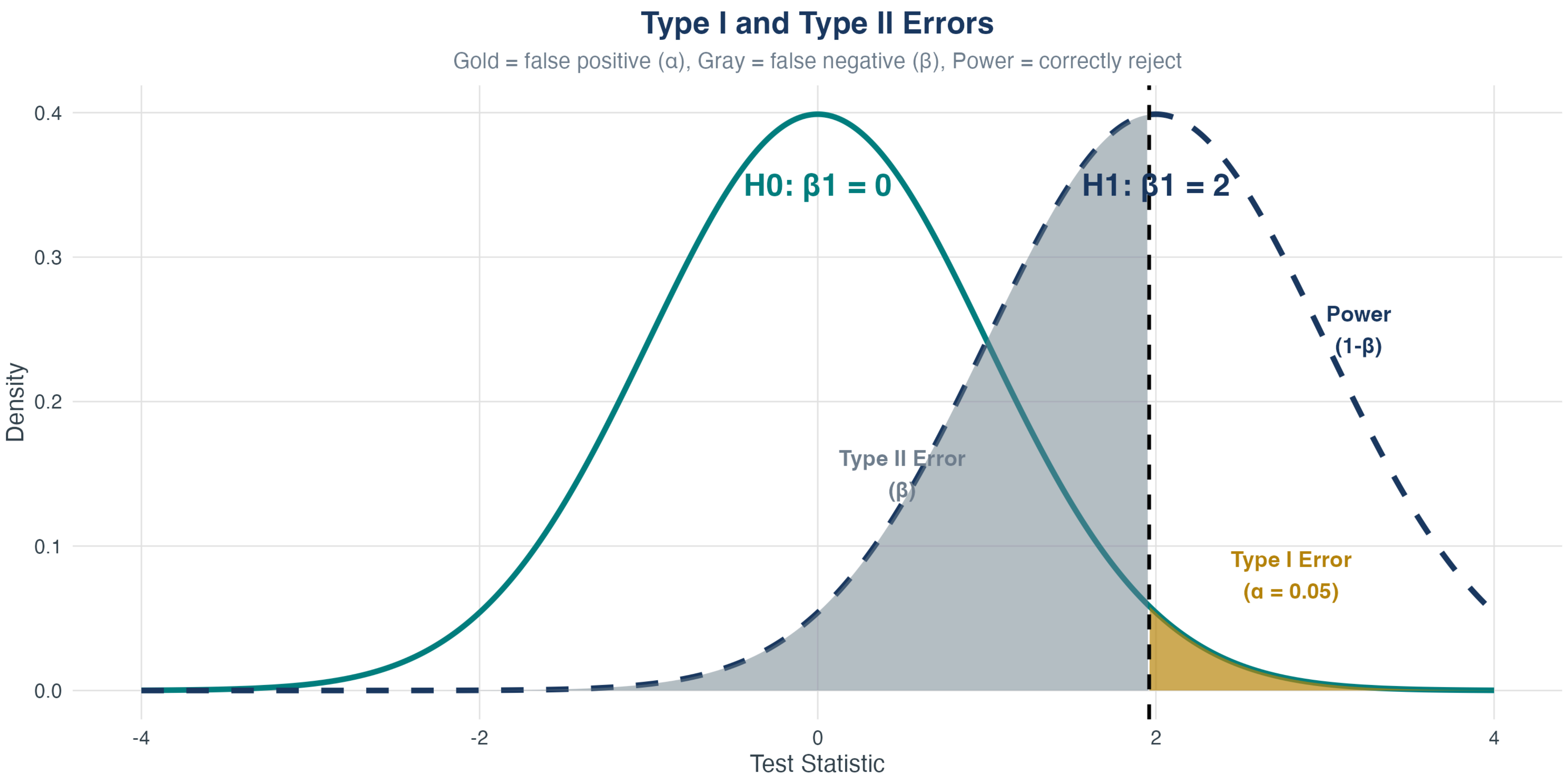
What is the t-statistic?

What is your decision?

Answer

1. {.fragment} $t = \dfrac{3.5 - 0}{1.2} = 2.92$
2. {.fragment} Critical value: $c_{\alpha/2} = 1.96$
3. {.fragment} Since $|2.92| > 1.96$, **reject** H_0 at 5% level. {.fragment} Conclusion: strong evidence that $\beta_1 \neq 0$.

Type I and Type II Errors



Type I and Type II errors

Knowledge Check: Type I and Type II Errors

Consider the following scenarios. For each, decide whether it is a **Type I error**, a **Type II error**, or **neither**:

A new medication does **not** actually help patients, but your study concludes that it **does**. What type of error (if any) has occurred?

The new medication **does** help patients, but your study fails to find evidence and concludes that it **doesn't work**. What type of error is this?

In a wage experiment, there is **no true effect** of a job training program on wages, but your study finds a statistically significant increase. What error (if any) is this?

There **is** a real positive effect of the training program on wages, but your study finds no statistically significant difference. What type of error is this?

Answers

Choosing a Significance Level

Why $\alpha = 0.05$ is conventional

We tolerate a 5% chance of false positives

Balances Type I and Type II errors

Widely accepted standard in social sciences

The tradeoff

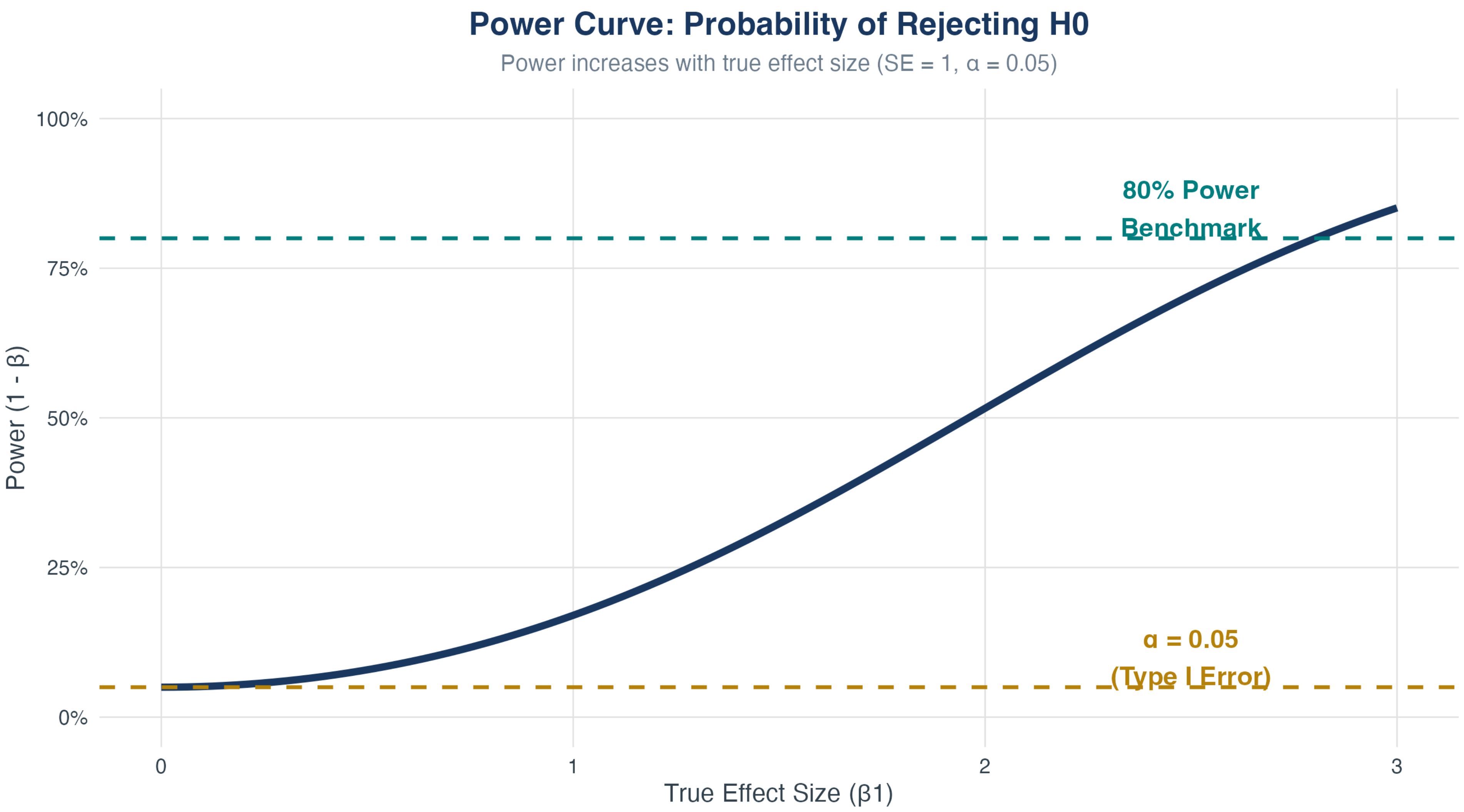
Smaller $\alpha \Rightarrow$ fewer false positives, but **more** false negatives (lower power)

Larger $\alpha \Rightarrow$ more false positives, but **fewer** false negatives (higher power)

You cannot minimize both at once.

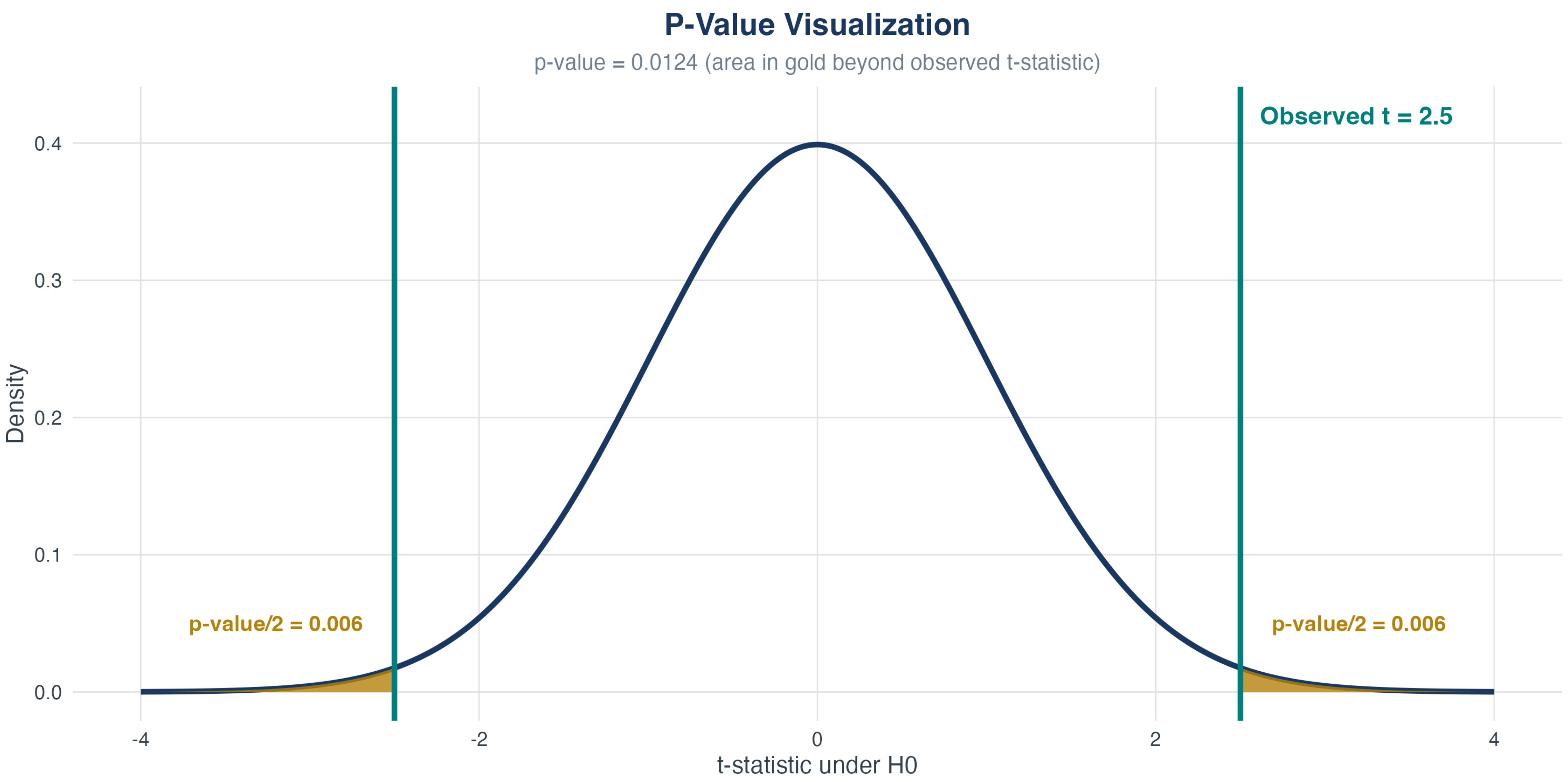
We typically fix α and try to increase power by increasing n.

Statistical Power



Power increases with:

Understanding p-Values



p-value interpretation

Hypothesis Testing in Stata

Testing $H_0 : \beta_1 = 0$

```
. regress bwght cigs

      Source |       SS           df       MS      Number of obs   =     1,388
-------------+----------------------------------   F(1, 1386)      =     32.24
       Model |  13060.4194         1  13060.4194   Prob > F        =    0.0000
    Residual |    561551.3     1,386  405.159668   R-squared       =    0.0227
-------------+----------------------------------   Adj R-squared   =    0.0220
       Total |   574611.72     1,387  414.283864   Root MSE        =    20.129

------------------------------------------------------------------------------
       bwght | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        cigs |  -.5137721   .0904909    -5.68   0.000    -.6912861   -.3362581
       _cons |   119.7719   .5723407   209.27   0.000     118.6492    120.8946
------------------------------------------------------------------------------
```

Stata output automatically shows:

Testing $H_0 : \beta_1 = a$ (Non-Zero Null)

What if we want to test a different null value?

Suppose we hypothesize that an extra year of education increases wage by exactly $3/hour:

$$H_0 : \beta_1 = 3 \quad \text{vs} \quad H_1 : \beta_1 \neq 3.$$

Modified t-statistic:

$$t = \frac{\hat{\beta}_1 - 3}{\text{SE}(\hat{\beta}_1)}.$$

Example

Economic vs Statistical Significance

⚠ **Critical distinction**

Statistical significance ≠ economic significance.

Scenario 1: Statistically significant but economically trivial

$\hat{\beta_1} = 0.05$, $\text{SE}(\hat{\beta_1}) = 0.01$, $t = 5$ (highly significant!)

An extra year of education increases wage by 5 cents/hour

Technically significant, but economically negligible

Scenario 2: Economically important but statistically insignificant

$\hat{\beta_1} = 8$, $\text{SE}(\hat{\beta_1}) = 10$, $t = 0.8$ (not significant)

Small sample ⇒ high uncertainty

Effect could be large, but we can't be confident

Conducting a Hypothesis Test

Is mother's education associated with birthweight?

```
. regress bwght motheduc

      Source |       SS           df       MS      Number of obs   =     1,387
-------------+----------------------------------   F(1, 1385)      =      6.65
       Model |  2745.15468         1  2745.15468   Prob > F        =    0.0100
    Residual |  571729.586     1,385  412.801145   R-squared       =    0.0048
-------------+----------------------------------   Adj R-squared   =    0.0041
       Total |  574474.741     1,386   414.48394   Root MSE        =    20.318

------------------------------------------------------------------------------
       bwght | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    motheduc |   .5921371   .2296198     2.58   0.010     .1416969    1.042577
       _cons |   111.0482   3.020007    36.77   0.000     105.1239    116.9725
------------------------------------------------------------------------------
```

Testing $H_0 : \beta_1 = a$

$$H_0 : \beta_1 = a$$

$$H_1 : \beta_1 \neq a$$

Adjust the t-statistic

$$t = \frac{\hat{\beta}_1 - a}{SE(\hat{\beta}_1)}$$

5.2 Confidence Intervals for β_1

From Hypothesis Tests to Confidence Intervals

Hypothesis test: Is β_1 equal to a specific value?

Confidence interval: What is a **plausible range** for β_1?

> ⓘ **95% Confidence Interval**
>
> A random interval that contains the true parameter 95% of the time in repeated samples.

Two equivalent interpretations:

The set of all null values we **cannot reject** at 5% level

An interval constructed from the data that **covers the truth** 95% of the time

Constructing a 95% Confidence Interval

Since for large samples,

$$t = \frac{\hat{\beta}_1 - \beta_1}{\mathrm{SE}(\hat{\beta}_1)} \sim N(0, 1)$$

$$P\left(-1.96 < \frac{\hat{\beta}_1 - \beta_1}{\mathrm{SE}(\hat{\beta}_1)} < 1.96\right) = 0.95.$$

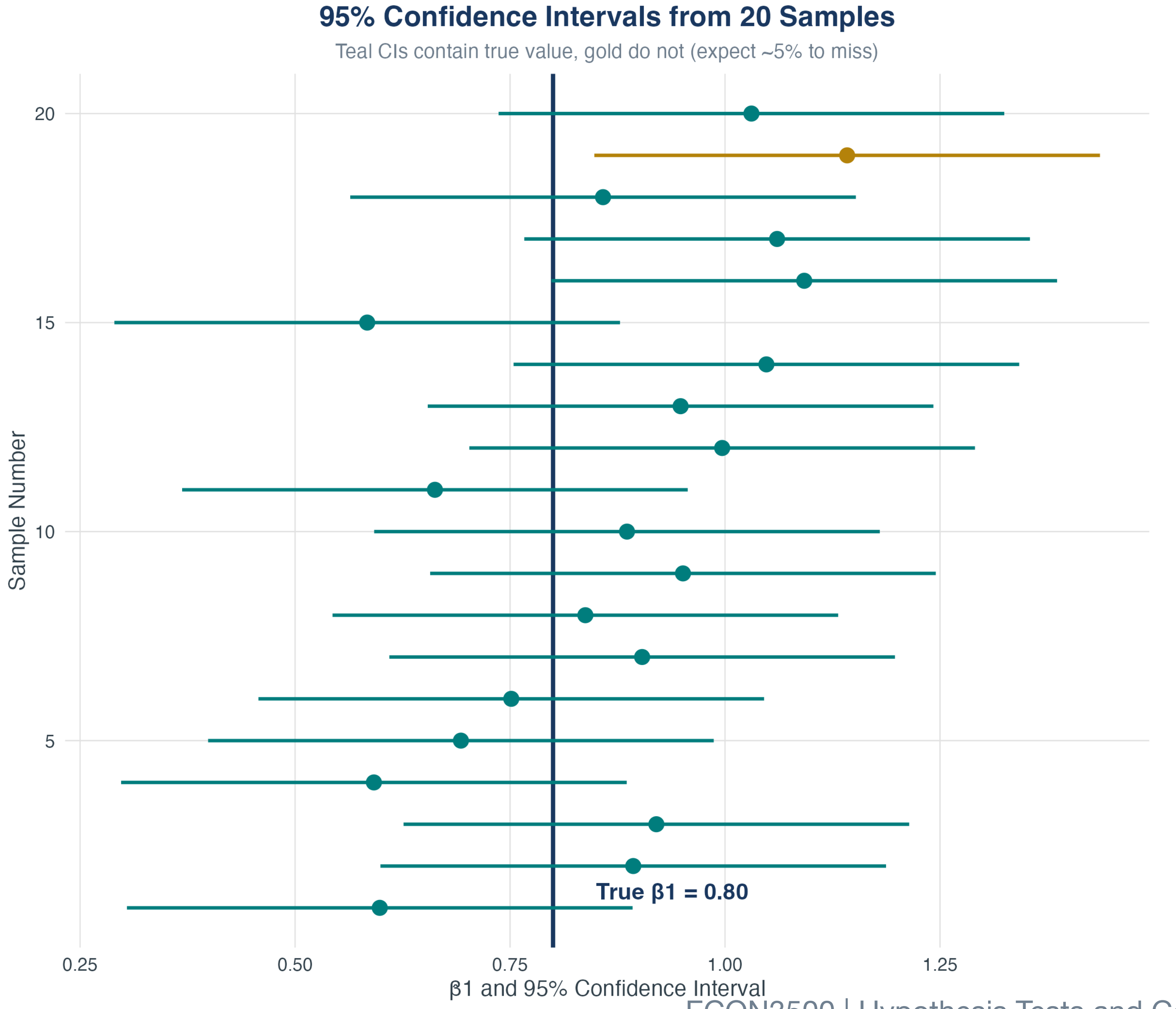
Rearranging,

$$P\left(\hat{\beta}_1 - 1.96 \cdot \mathrm{SE}(\hat{\beta}_1) < \beta_1 < \hat{\beta}_1 + 1.96 \cdot \mathrm{SE}(\hat{\beta}_1)\right) = 0.95.$$

95% Confidence Interval Formula

$$\hat{\beta}_1 \pm 1.96 \cdot \mathrm{SE}(\hat{\beta}_1).$$

Visualizing Confidence Intervals



Confidence Intervals for Other Confidence Levels

General formula

$$\hat{\beta}_1 \pm c_{\alpha/2} \cdot \mathrm{SE}(\hat{\beta}_1).$$

Common choices

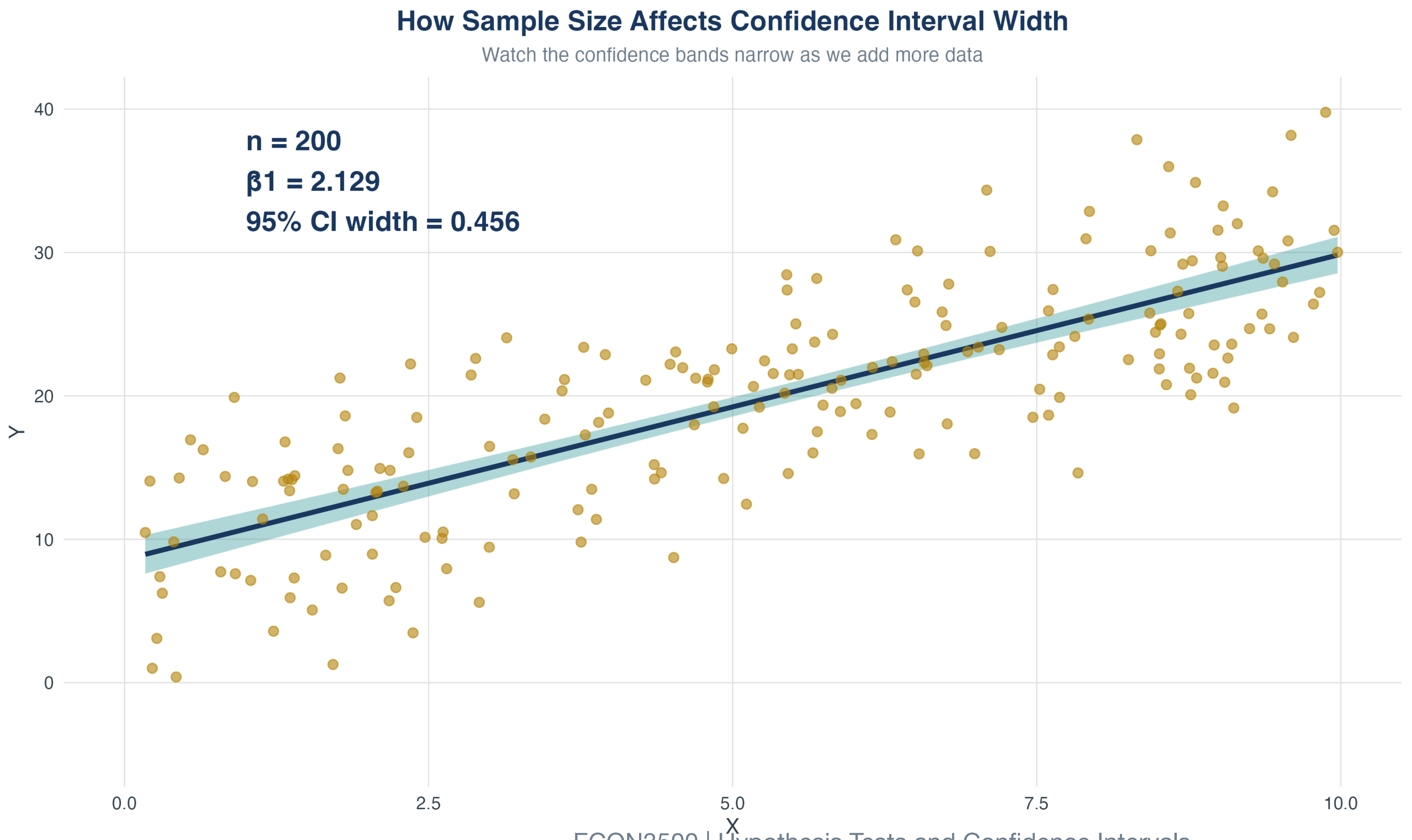
Confidence level		Critical value
90%	0.10	1.645
95%	0.05	1.96
99%	0.01	2.58

Tradeoff

Comparing Confidence Levels

How Sample Size Affects Precision

Animated: Adding Data Shrinks the CI



Confidence Intervals in Stata

```
. regress bwght motheduc

      Source |       SS           df       MS      Number of obs   =     1,387
-------------+----------------------------------   F(1, 1385)      =      6.65
       Model |  2745.15468         1  2745.15468   Prob > F        =    0.0100
    Residual |  571729.586     1,385  412.801145   R-squared       =    0.0048
-------------+----------------------------------   Adj R-squared   =    0.0041
       Total |  574474.741     1,386   414.48394   Root MSE        =    20.318

------------------------------------------------------------------------------
       bwght | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    motheduc |   .5921371   .2296198     2.58   0.010     .1416969    1.042577
       _cons |   111.0482   3.020007    36.77   0.000     105.1239    116.9725
------------------------------------------------------------------------------
```

Interpretation

The Link Between CIs and Hypothesis Tests

Key relationship

If $\beta_{1,0}$ is **inside** the 95% CI, we **fail to reject** $H_0 : \beta_1 = \beta_{1,0}$ at 5% level.

If $\beta_{1,0}$ is **outside** the 95% CI, we **reject** $H_0 : \beta_1 = \beta_{1,0}$ at 5% level.

Example

95% CI: $[1.8, 6.6]$

$H_0 : \beta_1 = 0$

- 0 not in CI $\Rightarrow$ reject

$H_0 : \beta_1 = 4$

- 4 is in CI $\Rightarrow$ fail to reject

5.3 Regression When X Is Binary

Regression When X Is Binary

Often, our regressor takes only two values:

$X = 1$ if small class, $X = 0$ if large class

$X = 1$ if treated, $X = 0$ if control

$X = 1$ if college degree, $X = 0$ if no degree

$X = 1$ if female, $X = 0$ if male

These are called **binary variables**, **dummy variables**, or **indicator variables**.

Terminology

$X = 1$: "treatment group" or "category of interest"

$X = 0$: "control group" or "reference category"

Gender is not binary, but it is recorded as binary in many datasets—data availability shapes our understanding of the world.

Interpreting a Binary X

Population model: $Y_i = \beta_0 + \beta_1 X_i + u_i$

When $X_i = 0$:

$Y_i = \beta_0 + \beta_1 * 0 + u_i$

$Y_i = \beta_0 + u_i$

$E[Y_i \mid X_i = 0] = E[\beta_0 + u_i] = E[\beta_0] + E[u_i] = \beta_0$

When $X_i = 1$:

$Y_i = \beta_0 + \beta_1 + u_i$

$Y_i = \beta_0 + \beta_1 * 1 + u_i$

$Y_i = \beta_0 + \beta_1 + u_i$

$E[Y_i \mid X_i = 1] = E[\beta_0 + \beta_1 + u_i] = E[\beta_0] + E[\beta_1] + E[u_i] = \beta_0 + \beta_1$

Interpreting a Binary X

Therefore:

Note

$$\beta_1 = E(Y_i \mid X_i = 1) - E(Y_i \mid X_i = 0)$$

So β_1 is the population difference in group means.

Interpreting a Binary X: Stata Output

Is sex associated with birthweight?

```
.    regress bwght male

      Source |       SS           df       MS      Number of obs   =     1,388
-------------+----------------------------------   F(1, 1386)      =      7.27
       Model |  2998.87965         1  2998.87965   Prob > F        =    0.0071
    Residual |   571612.84     1,386  412.419077   R-squared       =    0.0052
-------------+----------------------------------   Adj R-squared   =    0.0045
       Total |   574611.72     1,387  414.283864   Root MSE        =    20.308

------------------------------------------------------------------------------
       bwght | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        male |    2.94235   1.091149     2.70   0.007     .8018673    5.082832
       _cons |   117.1669   .7875145   148.78   0.000     115.6221    118.7118
------------------------------------------------------------------------------
```

Interpreting a Binary X (Example)

Is sex associated with birthweight?

$$\widehat{bwght} = 117.17 + 2.94\, male$$

Average birthweight of female babies: $E[bwght \mid male = 0] = 117.17$ ounces

Average birthweight of male babies: $E[bwght \mid male = 1] = 117.17 + 2.94 = 120.11$ ounces

5.4 Heteroskedasticity and Homoskedasticity

Heteroskedasticity and Homoskedasticity

What do the terms mean?

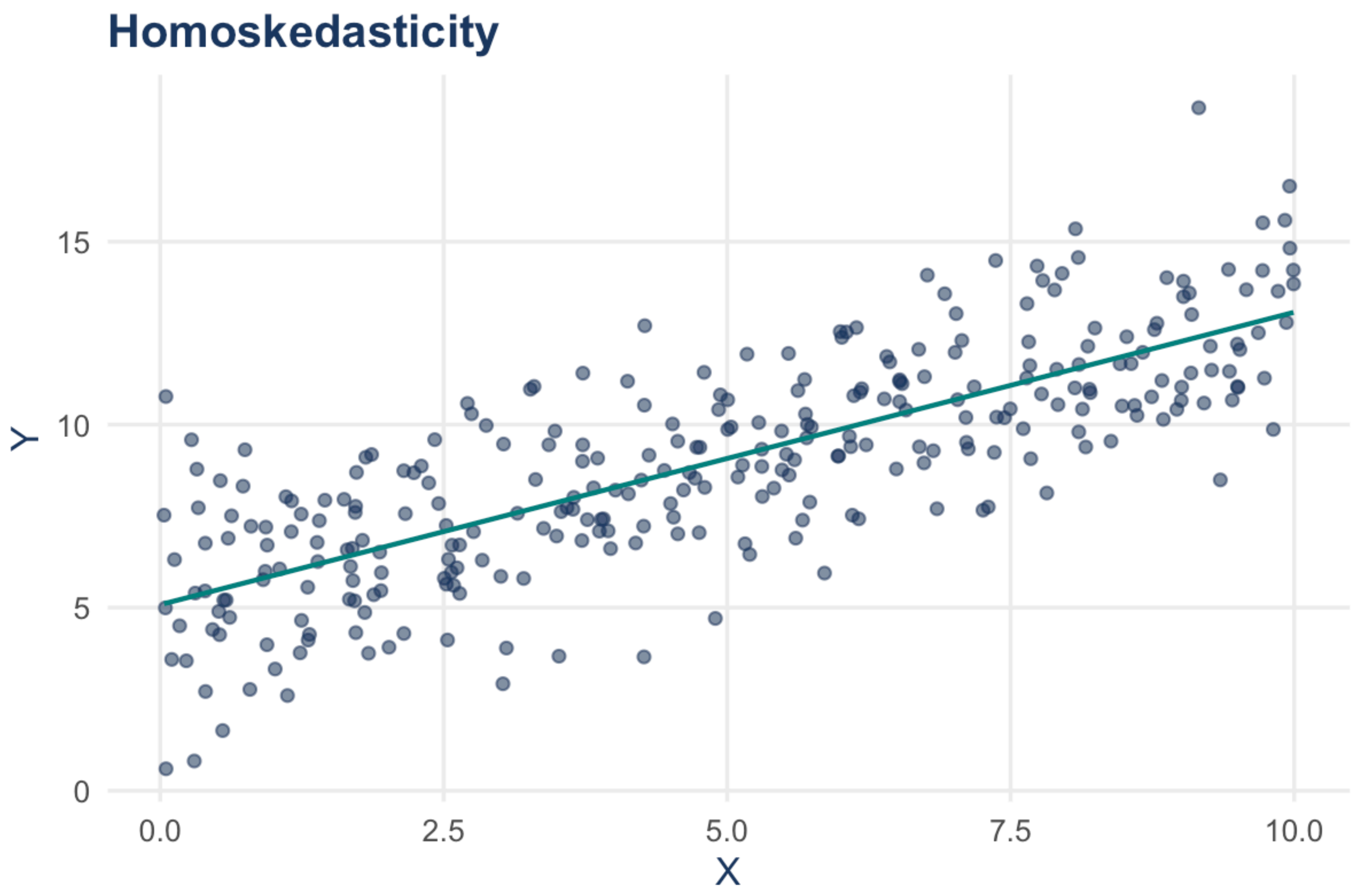
Consequences of heteroskedasticity

Implications for computing standard errors

If $var(u \mid X = x)$ is constant, u is **homoskedastic**.

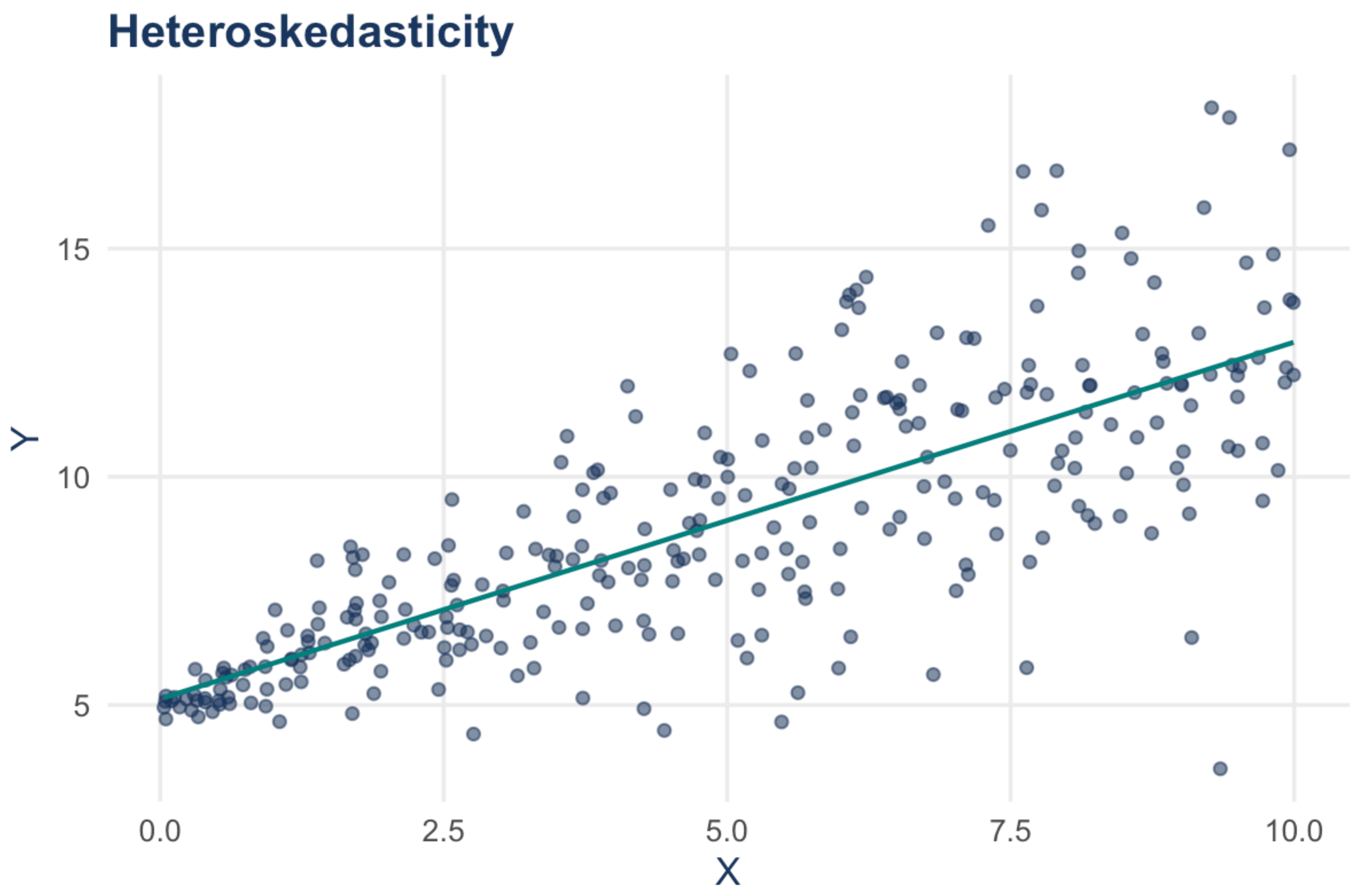
Otherwise, u is **heteroskedastic**.

Homoskedasticity in a Picture



The variance of u is constant

Heteroskedasticity in a Picture



The variance of u is not constant

Does Heteroskedasticity Affect $\hat{\beta}_1$?

Recall the least squares assumptions:

$E(u \mid X = x) = 0$

(X_i, Y_i) are i.i.d.

Large outliers are rare

Heteroskedasticity concerns $var(u \mid X = x)$.

Because we did **not** assume homoskedasticity, we have implicitly allowed heteroskedasticity.

So Who Cares?

Heteroskedasticity does **not** affect point estimates of β_1.

But it **does** affect your standard errors.

Homoskedastic-only standard errors are unbiased only under homoskedasticity. We adjust using **heteroskedasticity-robust** standard errors.

Heteroskedasticity-Robust Standard Errors (Stata)

```
. regress bwght cigs

      Source |       SS           df       MS      Number of obs   =     1,388
-------------+----------------------------------   F(1, 1386)      =     32.24
       Model |  13060.4194         1  13060.4194   Prob > F        =    0.0000
    Residual |    561551.3     1,386  405.159668   R-squared       =    0.0227
-------------+----------------------------------   Adj R-squared   =    0.0220
       Total |   574611.72     1,387  414.283864   Root MSE        =    20.129

------------------------------------------------------------------------------
       bwght | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        cigs |  -.5137721   .0904909    -5.68   0.000    -.6912861   -.3362581
       _cons |   119.7719   .5723407   209.27   0.000     118.6492    120.8946
------------------------------------------------------------------------------
```

Heteroskedasticity-Robust Standard Errors (Stata)

```
.  regress bwght cigs,robust

Linear regression                               Number of obs     =      1,388
                                                F(1, 1386)        =      34.29
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0227
                                                Root MSE          =     20.129

------------------------------------------------------------------------------
             |               Robust
       bwght | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        cigs |  -.5137721   .0877334    -5.86   0.000    -.6858767   -.3416675
       _cons |   119.7719   .5745494   208.46   0.000     118.6448     120.899
------------------------------------------------------------------------------
```

Comparing Homoskedastic and Robust SEs

Heteroskedasticity: The Bottom Line

Situation	Homoskedastic SE	Robust SE
Errors are homoskedastic	Correct	Correct
Errors are heteroskedastic	Wrong	Correct

Always use robust standard errors.

No downside if errors are homoskedastic, protects you if they're not.

5.5 Gauss–Markov Theorem

The Extended Least Squares Assumptions

Consider our three LS assumptions (needed for unbiasedness):

$E(u \mid X = x) = 0$

(X_i, Y_i) are i.i.d.

Large outliers are rare

Plus one more:

u is **homoskedastic**

Gauss–Markov Theorem

Under these four extended LS assumptions, $\hat{\beta}_1$ has the smallest variance among **all linear estimators**.

This is the **Gauss–Markov Theorem**.

Under GM, OLS estimators are **BLUE**:

Best

Linear

Unbiased

Estimators

OLS Limitations

Homoskedasticity often doesn't hold

GM is only for **linear** estimators (a small subset)

If we know the form of heteroskedasticity, we can use **weighted least squares**

With many outliers, **least absolute deviations (LAD)** can be more efficient

In most applied regression analysis, we use OLS—so that is what we will do, too.

Conclusion

Key ideas to remember:

Hypothesis tests and confidence intervals for β_1

Interpreting binary regressors

Why heteroskedasticity matters for standard errors

The Gauss–Markov conditions and what BLUE means