Introduction to Econometrics

Lecture 6 : OLS inference (SW Cha 5 & 7)

Zhaopeng Qu

Business School,Nanjing University

Oct. 18, 2018

Outlines

1. Review: Hypothesis Test
   - Hypothesis Test:
   - Simple OLS in Normal Sampling Distribution
2. OLS with One Regressor: Hypothesis Tests
   - Hypothesis Test of of $\bar{Y}$
   - OLS with One Regressor: Hypothesis Tests
   - Gauss-Markov theorem and Heteroskedasticity
3. OLS with Multiple Regressors: Hypotheses tests
   - Hypothesis test and Confidence interval for single coefficient

Review: Hypothesis Test

Hypothesis Testing

Definition

A hypothesis is a statement about a population parameter, thus θ. Formally, we want to test whether is significantly different from a certain value μ_0

$$H_0 : \theta = \mu_0$$

which is called **null hypothesis**. **The alternative hypothesis** is

$$H_1 : \theta \neq \mu_0$$

- If the value μ_0 does not lie within the calculated condence interval, then we **reject** the null hypothesis.
- If the value μ_0 lie within the calculated condence interval, then we **fail to reject** the null hypothesis.
- The two hypotheses must be disjoint: it should be the case that either H_0 is true or H_1 but never together simultaneously.

Hypothesis Testing

Definition

A hypothesis is a statement about a population parameter, thus θ. Formally, we want to test whether is significantly different from a certain value μ_0

$$H_0 : \theta = \mu_0$$

which is called **null hypothesis**. **The alternative hypothesis** is

$$H_1 : \theta \neq \mu_0$$

- If the value μ_0 does not lie within the calculated condence interval, then we **reject** the null hypothesis.
- If the value μ_0 lie within the calculated condence interval, then we **fail to reject** the null hypothesis.
- The two hypotheses must be disjoint: it should be the case that either H_0 is true or H_1 but never together simultaneously.

Hypothesis Testing

Definition

A hypothesis is a statement about a population parameter, thus θ. Formally, we want to test whether is significantly different from a certain value μ_0

$$H_0 : \theta = \mu_0$$

which is called **null hypothesis**. **The alternative hypothesis** is

$$H_1 : \theta \neq \mu_0$$

- If the value μ_0 does not lie within the calculated condence interval, then we **reject** the null hypothesis.
- If the value μ_0 lie within the calculated condence interval, then we **fail to reject** the null hypothesis.
- The two hypotheses must be disjoint: it should be the case that either H_0 is true or H_1 but never together simultaneously.

Hypothesis Testing

Definition

A hypothesis is a statement about a population parameter, thus θ. Formally, we want to test whether is significantly different from a certain value μ_0

$$H_0 : \theta = \mu_0$$

which is called **null hypothesis**. **The alternative hypothesis** is

$$H_1 : \theta \neq \mu_0$$

- If the value μ_0 does not lie within the calculated condence interval, then we **reject** the null hypothesis.
- If the value μ_0 lie within the calculated condence interval, then we **fail to reject** the null hypothesis.
- The two hypotheses must be disjoint: it should be the case that either H_0 is true or H_1 but never together simultaneously.

Two Type Errors

- A Type I error is when we *reject* the null hypothesis H_0 when it is in fact *true*.("left-wing"). The probability of Type I error is denoted by α and called **significance level** or size of a test.

$$P(Type\, I\, error) = P(reject\, H_0 \mid H_0 is\, true) = \alpha$$

- A Type II error is when we fail to reject the null hypothesis when it is false.("right-wing")

$$P(Type\, II\, error) = P(accept\, H_0 \mid H_0 is\, false)$$

- Unfortunately, the probabilities of Type I and II errors are inversely related. By decreasing the probability of Type I error α, one makes the critical region smaller, which increases the probability of the Type II error. Thus it is impossible to make both errors arbitrary small.

Two Type Errors

- A Type I error is when we *reject* the null hypothesis H_0 when it is in fact *true*.("left-wing"). The probability of Type I error is denoted by α and called **significance level** or size of a test.

$$P(Type\ I\ error) = P(reject\ H_0 \mid H_0 is\ true) = \alpha$$

- A Type II error is when we fail to reject the null hypothesis when it is false.("right-wing")

$$P(Type\ II\ error) = P(accept\ H_0 \mid H_0 is\ false)$$

- Unfortunately, the probabilities of Type I and II errors are inversely related. By decreasing the probability of Type I error α, one makes the critical region smaller, which increases the probability of the Type II error. Thus it is impossible to make both errors arbitrary small.

Two Type Errors

- A Type I error is when we *reject* the null hypothesis H_0 when it is in fact *true*.(“left-wing”). The probability of Type I error is denoted by α and called **significance level** or size of a test.

$$P(Type\ I\ error) = P(reject\ H_0 \mid H_0 is\ true) = \alpha$$

- A Type II error is when we fail to reject the null hypothesis when it is false.(“right-wing”)

$$P(Type\ II\ error) = P(accept\ H_0 \mid H_0 is\ false)$$

- Unfortunately, the probabilities of Type I and II errors are inversely related. By decreasing the probability of Type I error α, one makes the critical region smaller, which increases the probability of the Type II error. Thus it is impossible to make both errors arbitrary small.

Decision Rule

- Usually, we has to carry the "burden of proof," and the case that he is interested in is stated as H_1.
  - We would like to prove that his assertion H_1 is true by showing that the data rejects H_0.
- The decision rule that leads us to reject or not to reject H_0 is based on a test statistic, which is a function of the data

$$T_n = T(Y_1, ..., Y_n)$$

- Usually, one rejects H_0 if the test statistic falls into a **critical region**. A critical region is constructed by taking into account the probability of making a wrong decision.
- By convention, α is chosen to be a small number, for example, a $=$ 0.01, 0.05, or 0.10.

Decision Rule

- Usually, we has to carry the "burden of proof," and the case that he is interested in is stated as H_1.
  - We would like to prove that his assertion H_1 is true by showing that the data rejects H_0.
- The decision rule that leads us to reject or not to reject H_0 is based on a test statistic, which is a function of the data

$$T_n = T(Y_1, ..., Y_n)$$

- Usually, one rejects H_0 if the test statistic falls into a **critical region**. A critical region is constructed by taking into account the probability of making a wrong decision.
- By convention, α is chosen to be a small number, for example, a $=$ 0.01, 0.05, or 0.10.

Decision Rule

- Usually, we has to carry the "burden of proof," and the case that he is interested in is stated as H_1.
    - We would like to prove that his assertion H_1 is true by showing that the data rejects H_0.
- The decision rule that leads us to reject or not to reject H_0 is based on a test statistic, which is a function of the data

$$T_n = T(Y_1, ..., Y_n)$$

- Usually, one rejects H_0 if the test statistic falls into a **critical region**. A critical region is constructed by taking into account the probability of making a wrong decision.
- By convention, α is chosen to be a small number, for example, a $=$ 0.01, 0.05, or 0.10.

Decision Rule

- Usually, we has to carry the "burden of proof," and the case that he is interested in is stated as H_1.
  - We would like to prove that his assertion H_1 is true by showing that the data rejects H_0.
- The decision rule that leads us to reject or not to reject H_0 is based on a test statistic, which is a function of the data

$$T_n = T(Y_1, ..., Y_n)$$

- Usually, one rejects H_0 if the test statistic falls into a **critical region**. A critical region is constructed by taking into account the probability of making a wrong decision.
- By convention, α is chosen to be a small number, for example, $a = 0.01$, 0.05, or 0.10.

Decision Rule

- Usually, we has to carry the "burden of proof," and the case that he is interested in is stated as H_1.
  - We would like to prove that his assertion H_1 is true by showing that the data rejects H_0.
- The decision rule that leads us to reject or not to reject H_0 is based on a test statistic, which is a function of the data

$$T_n = T(Y_1, ..., Y_n)$$

- Usually, one rejects H_0 if the test statistic falls into a **critical region**. A critical region is constructed by taking into account the probability of making a wrong decision.
- By convention, α is chosen to be a small number, for example, a $=$ 0.01, 0.05, or 0.10.

Testing procedure

- The following are the steps of the hypothesis testing:
  1. Specify H_0 and H_1.
  2. Choose the significance level α.
  3. Define a decision rule (critical region).
  4. Given the data compute the test statistic and see if it falls into the critical region.

Testing procedure

- The following are the steps of the hypothesis testing:
  1. Specify H_0 and H_1.
  2. Choose the significance level α.
  3. Define a decision rule (critical region).
  4. Given the data compute the test statistic and see if it falls into the critical region.

Testing procedure

- The following are the steps of the hypothesis testing:
  1. Specify H_0 and H_1.
  2. Choose the significance level α.
  3. Define a decision rule (critical region).
  4. Given the data compute the test statistic and see if it falls into the critical region.

Testing procedure

- The following are the steps of the hypothesis testing:
  1. Specify H_0 and H_1.
  2. Choose the significance level α.
  3. Define a decision rule (critical region).
  4. Given the data compute the test statistic and see if it falls into the critical region.

Testing procedure

- The following are the steps of the hypothesis testing:
  1. Specify H_0 and H_1.
  2. Choose the significance level α.
  3. Define a decision rule (critical region).
  4. Given the data compute the test statistic and see if it falls into the critical region.

P-Value

- To provide additional information, we could ask the question: What is the *largest significance* level at which we could carry out the test and still fail to reject the null hypothesis?
- Or in other word, given the data, *the smallest significance* level at which the null can be rejected.
- We can consider the **p-value** of a test
  1. Calculate the t-statistic t
  2. The largest significance level at which we would fail to reject H_0 is the significance level associated with using t as our critical value

$$p - value = 1 - \Phi(t)$$

where $\Phi(t)$ denotes the standard normal c.d.f.(we assume that n is large enough)

P-Value

- To provide additional information, we could ask the question: What is the *largest significance* level at which we could carry out the test and still fail to reject the null hypothesis?
- Or in other word, given the data, *the smallest significance* level at which the null can be rejected.
- We can consider the **p-value** of a test
  1. Calculate the t-statistic t
  2. The largest significance level at which we would fail to reject H_0 is the significance level associated with using t as our critical value

$$p - value = 1 - \Phi(t)$$

where $\Phi(t)$ denotes the standard normal c.d.f.(we assume that n is large enough)

P-Value

- To provide additional information, we could ask the question: What is the *largest significance* level at which we could carry out the test and still fail to reject the null hypothesis?
- Or in other word, given the data, *the smallest significance* level at which the null can be rejected.
- We can consider the **p-value** of a test
  1. Calculate the t-statistic t
  2. The largest significance level at which we would fail to reject H_0 is the significance level associated with using t as our critical value

$$p - value = 1 - \Phi(t)$$

where $\Phi(t)$ denotes the standard normal c.d.f.(we assume that n is large enough)

P-Value

- To provide additional information, we could ask the question: What is the *largest significance* level at which we could carry out the test and still fail to reject the null hypothesis?
- Or in other word, given the data, *the smallest significance* level at which the null can be rejected.
- We can consider the **p-value** of a test
  1. Calculate the t-statistic t
  2. The largest significance level at which we would fail to reject H_0 is the significance level associated with using t as our critical value

$$p - value = 1 - \Phi(t)$$

where $\Phi(t)$ denotes the standard normal c.d.f.(we assume that n is large enough)

P-Value

- To provide additional information, we could ask the question: What is the *largest significance* level at which we could carry out the test and still fail to reject the null hypothesis?
- Or in other word, given the data, *the smallest significance* level at which the null can be rejected.
- We can consider the **p-value** of a test
  1. Calculate the t-statistic t
  2. The largest significance level at which we would fail to reject H_0 is the significance level associated with using t as our critical value

$$p - value = 1 - \Phi(t)$$

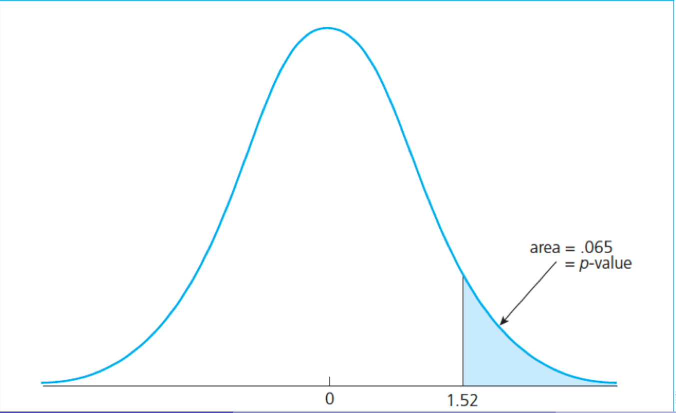
where $\Phi(t)$ denotes the standard normal c.d.f.(we assume that n is large enough)

P-Value: Case

- Suppose that $t = 1.52$, then we can find the largest significance level at which we would fail to reject H_0

$$p - value = P(T > 1.52 \mid H_0) = 1 - \Phi(1.52) = 0.065$$

FIGURE C.7 The p-value when $t = 1.52$ for the one-sided alternative $\mu > \mu_0$.



Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - **normal sampling distribution**

Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - **normal sampling distribution**

Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - **normal sampling distribution**

Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - **normal sampling distribution**

Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - **normal sampling distribution**

Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - **normal sampling distribution**

Statistical Inference

- Three Basic Assumption
  - Assumption 1
  - Assumption 2
  - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
  - unbiased
  - consistent
  - normal sampling distribution

Statistical Inference

- Three Basic Assumption
    - Assumption 1
    - Assumption 2
    - Assumption 3
- if the 3 least squares assumptions hold the OLS estimators
    - unbiased
    - consistent
    - **normal sampling distribution**

Sampling Distribution of β_1

- Recall: Sampling Distribution of $\bar{Y}$, based on the Central Limit theorem(C.L.T), the sample distribution in a large sample can approximates to a normal distribution.

$$\overline{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{n})$$

- So the sample distribution of β_1 in a large sample can also approximates to a normal distribution based on the Central Limit theorem(C.L.T), thus

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$$

- In last lecture We just showed you that

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{Var[(X_i - \mu_x)u_i]}{[Var(X_i)]^2}$$

  but did not prove that.
- Now we are going to derive it.

Sampling Distribution of β_1

- Recall: Sampling Distribution of $\bar{Y}$, based on the Central Limit theorem(C.L.T), the sample distribution in a large sample can approximates to a normal distribution.

$$\overline{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{n})$$

- So the sample distribution of β_1 in a large sample can also approximates to a normal distribution based on the Central Limit theorem(C.L.T), thus

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$$

- In last lecture We just showed you that

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{Var[(X_i - \mu_x)u_i]}{[Var(X_i)]^2}$$

but did not prove that.

- Now we are going to derive it.

Sampling Distribution of β_1

- Recall: Sampling Distribution of $\bar{Y}$, based on the Central Limit theorem(C.L.T), the sample distribution in a large sample can approximates to a normal distribution.

$$\overline{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{n})$$

- So the sample distribution of β_1 in a large sample can also approximates to a normal distribution based on the Central Limit theorem(C.L.T), thus

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$$

- In last lecture We just showed you that

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{Var[(X_i - \mu_x)u_i]}{[Var(X_i)]^2}$$

  but did not prove that.

- Now we are going to derive it.

Sampling Distribution of β_1

- Recall: Sampling Distribution of $\bar{Y}$, based on the Central Limit theorem(C.L.T), the sample distribution in a large sample can approximates to a normal distribution.

$$\overline{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{n})$$

- So the sample distribution of β_1 in a large sample can also approximates to a normal distribution based on the Central Limit theorem(C.L.T), thus

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$$

- In last lecture We just showed you that

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{Var[(X_i - \mu_x)u_i]}{[Var(X_i)]^2}$$

but did not prove that.

- Now we are going to derive it.

Sampling Distribution of β_1

- We derived that $\hat{\beta}_1$ in terms of regression and errors in following equation

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})}$$

- First consider the numerator of this term, thus $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})$
  - Because $\overline{X}$ is consistent, thus $\overline{X} \xrightarrow{p} \mu_x$.
  - And we know that $\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) = \sum_{i=1}^{n}(X_i - \overline{X})u_i$ and we let $v_i = (X_i - \mu_x)u_i$, then

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)u_i = \frac{1}{n}\sum_{i=1}^{n} v_i = \overline{v}$$

  - for Assumption 1, $E(v_i) = 0$, and Assumption 2, $\sigma_v^2 = Var[(X_i - \mu_x)u_i]$, then $\overline{v}$ is the sample mean of v_i, based on C.L.T,

$$\frac{\overline{v} - 0}{\sigma_{\overline{v}}} \xrightarrow{d} N(0,1) \; or \; \overline{v} \xrightarrow{d} N(0, \frac{\sigma_v^2}{n})$$

Sampling Distribution of β_1

- We derived that $\hat{\beta}_1$ in terms of regression and errors in following equation

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})}$$

- First consider the numerator of this term, thus $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})$
  - Because $\bar{X}$ is consistent, thus $X \xrightarrow{p} \mu_x$.
  - And we know that $\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) = \sum_{i=1}^{n}(X_i - \overline{X})u_i$ and we let $v_i = (X_i - \mu_x)u_i$, then

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)u_i = \frac{1}{n}\sum_{i=1}^{n} v_i = \bar{v}$$

  - for Assumption 1, $E(v_i) = 0$, and Assumption 2, $\sigma_v^2 = Var[(X_i - \mu_x)u_i]$, then $\bar{v}$ is the sample mean of v_i, based on C.L.T,

$$\frac{\bar{v} - 0}{\sigma_{\bar{v}}} \xrightarrow{d} N(0.1) \ or \ \bar{v} \xrightarrow{d} N(0, \frac{\sigma_v^2}{})$$

Sampling Distribution of β_1

- We derived that $\hat{\beta}_1$ in terms of regression and errors in following equation

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})}$$

- First consider the numerator of this term, thus $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})$
  - Because $\bar{X}$ is consistent, thus $X \xrightarrow{p} \mu_x$.
  - And we know that $\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) = \sum_{i=1}^{n}(X_i - \overline{X})u_i$ and we let $v_i = (X_i - \mu_x)u_i$, then

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)u_i = \frac{1}{n}\sum_{i=1}^{n} v_i = \bar{v}$$

  - for Assumption 1, $E(v_i) = 0$, and Assumption 2, $\sigma_v^2 = Var[(X_i - \mu_x)u_i]$, then $\bar{v}$ is the sample mean of v_i, based on C.L.T,

$$\frac{\bar{v} - 0}{\sigma_{\bar{v}}} \xrightarrow{d} N(0.1) \ or \ \bar{v} \xrightarrow{d} N(0, \frac{\sigma_v^2}{n})$$

Sampling Distribution of β_1

- We derived that $\hat{\beta}_1$ in terms of regression and errors in following equation

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})}$$

- First consider the numerator of this term, thus $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})$
  - Because $\bar{X}$ is consistent, thus $X \xrightarrow{p} \mu_x$.
  - And we know that $\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) = \sum_{i=1}^{n}(X_i - \overline{X})u_i$ and we let $v_i = (X_i - \mu_x)u_i$, then

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)u_i = \frac{1}{n}\sum_{i=1}^{n} v_i = \bar{v}$$

  - for Assumption 1, $E(v_i) = 0$, and Assumption 2, $\sigma_v^2 = Var[(X_i - \mu_x)u_i]$, then $\bar{v}$ is the sample mean of v_i, based on C.L.T,

$$\frac{\bar{v} - 0}{\sigma_{\bar{v}}} \xrightarrow{d} N(0.1) \ or \ \bar{v} \xrightarrow{d} N(0, \frac{\sigma_v^2}{})$$

Sampling Distribution of β_1

- We derived that $\hat{\beta}_1$ in terms of regression and errors in following equation

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})}$$

- First consider the numerator of this term, thus $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u})$
  - Because $\bar{X}$ is consistent, thus $X \xrightarrow{p} \mu_x$.
  - And we know that $\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) = \sum_{i=1}^{n}(X_i - \overline{X})u_i$ and we let $v_i = (X_i - \mu_x)u_i$, then

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}) \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)u_i = \frac{1}{n}\sum_{i=1}^{n} v_i = \bar{v}$$

  - for Assumption 1, $E(v_i) = 0$, and Assumption 2, $\sigma_v^2 = Var[(X_i - \mu_x)u_i]$, then $\bar{v}$ is the sample mean of v_i, based on C.L.T,

$$\frac{\bar{v} - 0}{\sigma_{\bar{v}}} \xrightarrow{d} N(0.1)\ or\ \bar{v} \xrightarrow{d} N(0, \frac{\sigma_v^2}{n})$$

Sampling Distribution of β_1

- Next consider the expression in the denominator, $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})$
  - this is the sample variance of X(except dividing by n rather than $n-1$, which is inconsequential if n is large)
  - As discussed in Section 3.2 [Equation (3.8)], the sample variance is a consistent estimator of the population variance.
- Combining these two results, we have that, in large samples

$$\hat{\beta}_1 - \beta_1 \cong \frac{\bar{v}}{Var[X_i]}$$

- Based on the characteristics of Normal distribution, then $\frac{\bar{v}}{Var[X_i]} \xrightarrow{d} N\left(0, \frac{\sigma_v^2}{n[Var(X_i)]^2}\right)$
- So $\hat{\beta}_1 \xrightarrow{d} N(\beta_1, \sigma^2_{\hat{\beta}_1})$ where $\sigma^2_{\hat{\beta}_1} = \frac{\sigma^2_{v_i}}{n[Var(X_i)]^2} = \frac{Var[(X_i-\mu_x)u_i]}{n[Var(X_i)]^2}$.

Sampling Distribution of β_1

- Next consider the expression in the denominator, $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(X_i-\overline{X})$
  - this is the sample variance of X(except dividing by n rather than $n-1$, which is inconsequential if n is large)
  - As discussed in Section 3.2 [Equation (3.8)], the sample variance is a consistent estimator of the population variance.
- Combining these two results, we have that, in large samples

$$\hat{\beta}_1-\beta_1 \cong \frac{\bar{v}}{Var[X_i]}$$

- Based on the characteristics of Normal distribution, then $\frac{\bar{v}}{Var[X_i]} \xrightarrow{d} N\left(0, \frac{\sigma_v^2}{n[Var(X_i)]^2}\right)$
- So $\hat{\beta}_1 \xrightarrow{d} N(\beta_1, \sigma^2_{\hat{\beta}_1})$ where $\sigma^2_{\hat{\beta}_1}=\frac{\sigma^2_{v_i}}{n[Var(X_i)]^2}=\frac{Var[(X_i-\mu_x)u_i]}{n[Var(X_i)]^2}$.

Sampling Distribution of β_1

- Next consider the expression in the denominator, $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(X_i-\overline{X})$
  - this is the sample variance of X(except dividing by n rather than $n-1$, which is inconsequential if n is large)
  - As discussed in Section 3.2 [Equation (3.8)], the sample variance is a consistent estimator of the population variance.
- Combining these two results, we have that, in large samples

$$\hat{\beta}_1-\beta_1 \cong \frac{\bar{v}}{Var[X_i]}$$

- Based on the characteristics of Normal distribution, then $\frac{\bar{v}}{Var[X_i]} \xrightarrow{d} N\left(0, \frac{\sigma_v^2}{n[Var(X_i)]^2}\right)$
- So $\hat{\beta}_1 \xrightarrow{d} N(\beta_1, \sigma^2_{\hat{\beta}_1})$ where $\sigma^2_{\hat{\beta}_1} = \frac{\sigma^2_{v_i}}{n[Var(X_i)]^2} = \frac{Var[(X_i-\mu_x)u_i]}{n[Var(X_i)]^2}$.

Sampling Distribution of β_1

- Next consider the expression in the denominator, $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(X_i-\overline{X})$
  - this is the sample variance of X(except dividing by n rather than $n-1$, which is inconsequential if n is large)
  - As discussed in Section 3.2 [Equation (3.8)], the sample variance is a consistent estimator of the population variance.
- Combining these two results, we have that, in large samples

$$\hat{\beta}_1-\beta_1 \cong \frac{\bar{v}}{Var[X_i]}$$

- Based on the characteristics of Normal distribution, then $\frac{\bar{v}}{Var[X_i]} \xrightarrow{d} N\left(0, \frac{\sigma_v^2}{n[Var(X_i)]^2}\right)$
- So $\hat{\beta}_1 \xrightarrow{d} N(\beta_1, \sigma^2_{\hat{\beta}_1})$ where $\sigma^2_{\hat{\beta}_1} = \frac{\sigma^2_{v_i}}{n[Var(X_i)]^2} = \frac{Var[(X_i-\mu_x)u_i]}{n[Var(X_i)]^2}$.

Sampling Distribution of β_1

- Next consider the expression in the denominator, $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(X_i-\overline{X})$
  - this is the sample variance of X(except dividing by n rather than $n-1$, which is inconsequential if n is large)
  - As discussed in Section 3.2 [Equation (3.8)], the sample variance is a consistent estimator of the population variance.
- Combining these two results, we have that, in large samples

$$\hat{\beta}_1-\beta_1 \cong \frac{\bar{v}}{Var[X_i]}$$

- Based on the characteristics of Normal distribution, then $\frac{\bar{v}}{Var[X_i]} \xrightarrow{d} N\left(0, \frac{\sigma_v^2}{n[Var(X_i)]^2}\right)$
- So $\hat{\beta}_1 \xrightarrow{d} N(\beta_1, \sigma^2_{\hat{\beta}_1})$ where $\sigma^2_{\hat{\beta}_1} = \frac{\sigma^2_{v_i}}{n[Var(X_i)]^2} = \frac{Var[(X_i-\mu_x)u_i]}{n[Var(X_i)]^2}$.

Sampling Distribution of β_1

- Next consider the expression in the denominator, $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(X_i-\overline{X})$
  - this is the sample variance of X(except dividing by n rather than $n-1$, which is inconsequential if n is large)
  - As discussed in Section 3.2 [Equation (3.8)], the sample variance is a consistent estimator of the population variance.
- Combining these two results, we have that, in large samples

$$\hat{\beta}_1-\beta_1 \cong \frac{\bar{v}}{Var[X_i]}$$

- Based on the characteristics of Normal distribution, then $\frac{\bar{v}}{Var[X_i]} \xrightarrow{d} N\left(0, \frac{\sigma_v^2}{n[Var(X_i)]^2}\right)$
- So $\hat{\beta}_1 \xrightarrow{d} N(\beta_1, \sigma^2_{\hat{\beta}_1})$ where $\sigma^2_{\hat{\beta}_1}=\frac{\sigma^2_{v_i}}{n[Var(X_i)]^2}=\frac{Var[(X_i-\mu_x)u_i]}{n[Var(X_i)]^2}$.

OLS with One Regressor: Hypothesis Tests

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\overline{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $|t^{act}| > critical\ value$
    - or if $p - value < significance\ level$

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\overline{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $|t^{act}| > critical\ value$
    - or if $p-value < significance\ level$

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\overline{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $|t^{act}| > critical\ value$
    - or if $p - value < significance\ level$

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\overline{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $|t^{act}| > critical\ value$
    - or if $p - value < significance\ level$

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\bar{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\overline{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Statistical Inference of $\bar{Y}$

- $H_0 : E[Y] = \mu_{Y,0}$ $H_1 : E[Y] \neq \mu_{Y,0}$
  - Step1: Compute the sample average $\bar{Y}$
  - Step2: Compute the **standard error** of $\bar{Y}$

$$SE(\overline{Y}) = \frac{s_Y}{\sqrt{n}}$$

  - Step3: Compute the t-statistic

$$t^{act} = \frac{\bar{Y} - \mu_{Y,0}}{SE\left(\bar{Y}\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Testing a 2-sided hypothesis concerning β_1

- Testing procedure for the population mean is justified by the Central Limit theorem.
- Central Limit theorem states that the t-statistic (standardized sample average) has an approximate $N(0, 1)$ distribution in large samples.
- Central Limit Theorem also states that
  - $\hat{\beta}_0$ & $\hat{\beta}_1$ have an approximate normal distribution in large samples.
  - and the standardized regression coefficients have approximate $N(0, 1)$ distribution in large samples.
- We can therefore use same general approach to test hypotheses about $\hat{\beta}_0$ & $\hat{\beta}_1$ under the Least Squares assumptions holding!

Testing a 2-sided hypothesis concerning β_1

- Testing procedure for the population mean is justified by the Central Limit theorem.
- Central Limit theorem states that the t-statistic (standardized sample average) has an approximate $N(0,1)$ distribution in large samples.
- Central Limit Theorem also states that
  - $\hat{\beta}_0$ & $\hat{\beta}_1$ have an approximate normal distribution in large samples.
  - and the standardized regression coefficients have approximate $N(0,1)$ distribution in large samples.
- We can therefore use same general approach to test hypotheses about $\hat{\beta}_0$ & $\hat{\beta}_1$ under the Least Squares assumptions holding!

Testing a 2-sided hypothesis concerning β_1

- Testing procedure for the population mean is justified by the Central Limit theorem.
- Central Limit theorem states that the t-statistic (standardized sample average) has an approximate $N(0,1)$ distribution in large samples.
- Central Limit Theorem also states that
  - $\hat{\beta}_0$ & $\hat{\beta}_1$ have an approximate normal distribution in large samples.
  - and the standardized regression coefficients have approximate $N(0,1)$ distribution in large samples.
- We can therefore use same general approach to test hypotheses about $\hat{\beta}_0$ & $\hat{\beta}_1$ under the Least Squares assumptions holding!

Testing a 2-sided hypothesis concerning β_1

- Testing procedure for the population mean is justified by the Central Limit theorem.
- Central Limit theorem states that the t-statistic (standardized sample average) has an approximate $N(0,1)$ distribution in large samples.
- Central Limit Theorem also states that
  - $\hat{\beta}_0$ & $\hat{\beta}_1$ have an approximate normal distribution in large samples.
  - and the standardized regression coefficients have approximate $N(0,1)$ distribution in large samples.
- We can therefore use same general approach to test hypotheses about $\hat{\beta}_0$ & $\hat{\beta}_1$ under the Least Squares assumptions holding!

Testing a 2-sided hypothesis concerning β_1

- Testing procedure for the population mean is justified by the Central Limit theorem.
- Central Limit theorem states that the t-statistic (standardized sample average) has an approximate $N(0,1)$ distribution in large samples.
- Central Limit Theorem also states that
  - $\hat{\beta}_0$ & $\hat{\beta}_1$ have an approximate normal distribution in large samples.
  - and the standardized regression coefficients have approximate $N(0,1)$ distribution in large samples.
- We can therefore use same general approach to test hypotheses about $\hat{\beta}_0$ & $\hat{\beta}_1$ under the Least Squares assumptions holding!

Testing a 2-sided hypothesis concerning β_1

- Testing procedure for the population mean is justified by the Central Limit theorem.
- Central Limit theorem states that the t-statistic (standardized sample average) has an approximate $N(0, 1)$ distribution in large samples.
- Central Limit Theorem also states that
  - $\hat{\beta}_0$ & $\hat{\beta}_1$ have an approximate normal distribution in large samples.
  - and the standardized regression coefficients have approximate $N(0, 1)$ distribution in large samples.
- We can therefore use same general approach to test hypotheses about $\hat{\beta}_0$ & $\hat{\beta}_1$ under the Least Squares assumptions holding!

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \; H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \ H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \ H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \ H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

OLS with One Regressor: Hypothesis Tests

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \; H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \; H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = \beta \; H_1 : \beta_1 \neq \beta$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_i + u_i$ by OLS to obtain $\hat{\beta}_1$
  - Step2: Compute the **standard error** of $\hat{\beta}_1$
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

OLS with One Regressor: Hypothesis Tests

The Standard error of $\hat{\beta}_1$

- The **standard error** of $\hat{\beta}_1$ is an estimator of the standard deviation of the sampling distribution $\sigma_{\hat{\beta}_1}$
- Recall from the last class

$$\sigma_{\hat{\beta}_1} = \sqrt{\frac{1}{n}\frac{Var[(X_i - \mu_X)\mu_i]}{[Var(X_i)]^2}}$$

  - We use sample variance $\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2$ to estimate population covariance $Var[(X_i - \mu_X)u_i]$
  - We also use $\frac{1}{n}\sum(X_i - \bar{X})^2$ to replace population covariance $Var(X_i)$
  - Then it can be shown that

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

The Standard error of $\hat{\beta}_1$

- The **standard error** of $\hat{\beta}_1$ is an estimator of the standard deviation of the sampling distribution $\sigma_{\hat{\beta}_1}$
- Recall from the last class

$$\sigma_{\hat{\beta}_1} = \sqrt{\frac{1}{n}\frac{Var[(X_i - \mu_X)\mu_i]}{[Var(X_i)]^2}}$$

- We use sample variance $\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2$ to estimate population covariance $Var[(X_i - \mu_X)u_i]$
- We also use $\frac{1}{n}\sum(X_i - \bar{X})^2$ to replace population covariance $Var(X_i)$
- Then it can be shown that

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

The Standard error of $\hat{\beta}_1$

- The **standard error** of $\hat{\beta}_1$ is an estimator of the standard deviation of the sampling distribution $\sigma_{\hat{\beta}_1}$
- Recall from the last class

$$\sigma_{\hat{\beta}_1} = \sqrt{\frac{1}{n}\frac{Var[(X_i - \mu_X)\mu_i]}{[Var(X_i)]^2}}$$

  - We use sample variance $\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2$ to estimate population covariance $Var[(X_i - \mu_X)u_i]$
  - We also use $\frac{1}{n}\sum(X_i - \bar{X})^2$ to replace population covariance $Var(X_i)$
  - Then it can be shown that

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

The Standard error of $\hat{\beta}_1$

- The **standard error** of $\hat{\beta}_1$ is an estimator of the standard deviation of the sampling distribution $\sigma_{\hat{\beta}_1}$
- Recall from the last class

$$\sigma_{\hat{\beta}_1} = \sqrt{\frac{1}{n}\frac{Var[(X_i - \mu_X)\mu_i]}{[Var(X_i)]^2}}$$

  - We use sample variance $\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2$ to estimate population covariance $Var[(X_i - \mu_X)u_i]$
  - We also use $\frac{1}{n}\sum(X_i - \bar{X})^2$ to replace population covariance $Var(X_i)$
  - Then it can be shown that

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

The Standard error of $\hat{\beta}_1$

- The **standard error** of $\hat{\beta}_1$ is an estimator of the standard deviation of the sampling distribution $\sigma_{\hat{\beta}_1}$
- Recall from the last class

$$\sigma_{\hat{\beta}_1} = \sqrt{\frac{1}{n}\frac{Var[(X_i - \mu_X)\mu_i]}{[Var(X_i)]^2}}$$

  - We use sample variance $\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2$ to estimate population covariance $Var[(X_i - \mu_X)u_i]$
  - We also use $\frac{1}{n}\sum(X_i - \bar{X})^2$ to replace population covariance $Var(X_i)$
  - Then it can be shown that

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2\hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

Testing a 2-sided hypothesis concerning β_1

- The simple OLS regression : $TestScore_i = \beta_0 + \beta_1 ClassSize_i + u_i$
- We run it in Stata

```
. regress test_score class_size, robust

Linear regression                               Number of obs     =        420
                                                  F(1, 418)         =      19.26
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.0512
                                                  Root MSE          =     18.581
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-2.279808	.5194892	-4.39	0.000	-3.300945	-1.258671
_cons	698.933	10.36436	67.44	0.000	678.5602	719.3057

Testing a 2-sided hypothesis concerning β_1

- The simple OLS regression : $TestScore_i = \beta_0 + \beta_1 ClassSize_i + u_i$
- We run it in Stata

```
. regress test_score class_size, robust

Linear regression                               Number of obs     =        420
                                                  F(1, 418)         =      19.26
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.0512
                                                  Root MSE          =     18.581

------------------------------------------------------------------------------
             |               Robust
  test_score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  class_size |  -2.279808   .5194892    -4.39   0.000    -3.300945   -1.258671
       _cons |    698.933   10.36436    67.44   0.000     678.5602    719.3057
------------------------------------------------------------------------------
```

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value\ 1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $| t^{act} | = | -4.39 | > critical\ value.1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value. 1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

 OLS with One Regressor: Hypothesis Tests

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value, 1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value.1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\, value.1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

Testing a 2-sided hypothesis concerning β_1

- $H_0 : \beta_1 = 0 \; H_1 : \beta_1 \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

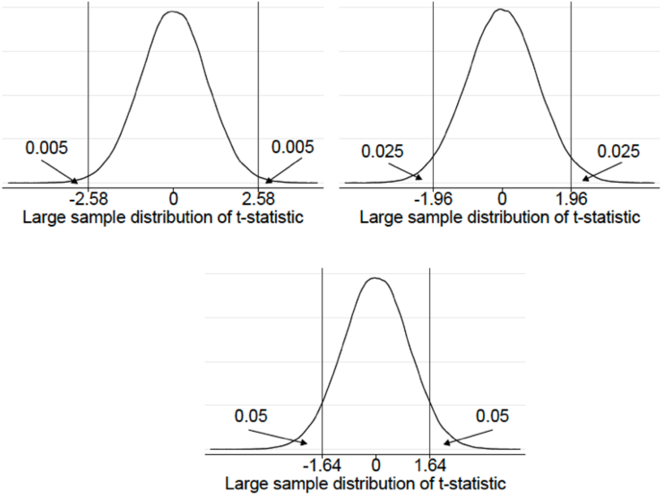
$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - 0}{0.52} = -4.39$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value.1.96$
  - $p - value = 0.00 < significance\ level = 0.05$

 OLS with One Regressor: Hypothesis Tests

Critical value of the t-statistic

The critical value of t-statistic depends on significance level α



-

1% and 10% significant levels

- Step 4: We reject the null hypothesis at a **10%** significance level because
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value.1.64$
  - $p - value = 0.00 < significance\ level = 0.1$
- Step 4: We reject the null hypothesis at a **1%** significance level because $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value.2.58$
  - $p - value = 0.00 < significance\ level = 0.01$

1% and 10% significant levels

- Step 4: We reject the null hypothesis at a **10%** significance level because
  - $| t^{act} | = | -4.39 | > critical\ value.1.64$
  - $p - value = 0.00 < significance\ level = 0.1$
- Step 4: We reject the null hypothesis at a **1%** significance level because
  $| t^{act} | = | -4.39 | > critical\ value.2.58$
  - $p - value = 0.00 < significance\ level = 0.01$

1% and 10% significant levels

- Step 4: We reject the null hypothesis at a **10%** significance level because
  - $| t^{act} |=| -4.39 |> critical\ value.1.64$
  - $p - value = 0.00 < significance\ level = 0.1$
- Step 4: We reject the null hypothesis at a **1%** significance level because
  $| t^{act} |=| -4.39 |> critical\ value.2.58$
  - $p - value = 0.00 < significance\ level = 0.01$

1% and 10% significant levels

- Step 4: We reject the null hypothesis at a **10%** significance level because
  - $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value.1.64$
  - $p - value = 0.00 < significance\ level = 0.1$
- Step 4: We reject the null hypothesis at a **1%** significance level because
  $\mid t^{act} \mid = \mid -4.39 \mid > critical\ value.2.58$
  - $p - value = 0.00 < significance\ level = 0.01$

1% and 10% significant levels

- Step 4: We reject the null hypothesis at a **10%** significance level because
  - $| t^{act} |=| -4.39 |> critical\ value.1.64$
  - $p - value = 0.00 < significance\ level = 0.1$
- Step 4: We reject the null hypothesis at a **1%** significance level because
  $| t^{act} |=| -4.39 |> critical\ value.2.58$
  - $p - value = 0.00 < significance\ level = 0.01$

Testing a 2-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 \neq -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $\mid t^{act} \mid = \mid -0.54 \mid < critical\ value\ 1.96$

Testing a 2-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 \neq -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $\mid t^{act} \mid = \mid -0.54 \mid < critical\ value.1.96$

Testing a 2-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 \neq -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $\mid t^{act} \mid = \mid -0.54 \mid < critical\ value.1.96$

 OLS with One Regressor: Hypothesis Tests

Testing a 2-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 \neq -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $| t^{act} | = | -0.54 | < critical\ value\ 1.96$

Testing a 2-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2$ $H_1 : \beta_1 \neq -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $\mid t^{act} \mid = \mid -0.54 \mid < critical\ value.1.96$

Testing a 2-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 \neq -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $\mid t^{act} \mid = \mid -0.54 \mid < critical\ value.1.96$

Testing a 1-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \ H_1 : \beta_1 < -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $t^{act} = -0.54 > critical\ value. -1.96$

Testing a 1-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \ H_1 : \beta_1 < -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $t^{act} = -0.54 > critical\ value. -1.96$

Testing a 1-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 < -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $t^{act} = -0.54 > critical\ value. -1.96$

Testing a 1-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \ H_1 : \beta_1 < -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
  - $t^{act} = -0.54 > critical\ value. -1.96$

Testing a 1-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 < -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
    - $t^{act} = -0.54 > critical\ value. -1.96$

Testing a 1-sided hypothesis concerning $\beta_1 = \beta$

- $H_0 : \beta_1 = -2 \; H_1 : \beta_1 < -2$
- Step1: Estimate $\hat{\beta}_1 = -2.28$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.52$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-2.28 - (-2)}{0.52} = -0.54$$

- Step4: we can't reject the null hypothesis at 5% significant level because
    - $t^{act} = -0.54 > critical\ value. -1.96$

Confidence interval for a regression coefficient β_1

- Method for constructing a confidence interval for a population mean can be easily extended to constructing a confidence interval for a regression coefficient.
- Using a two-sided test, a hypothesized value for β_1 will be rejected at *5% significance level* if $| t^{act} | > critical\ value.1.96.$
- So and will be in the confidence set if $| t^{act} | \leq critical\ value.1.96.$
- Thus the 95% confidence interval for β_1 are within ± 1.96 standard errors of $\hat{\beta}_1$

$$\hat{\beta}_1 \pm 1.96 \cdot SE\left(\hat{\beta}_1\right)$$

Confidence interval for a regression coefficientβ_1

- Method for constructing a confidence interval for a population mean can be easily extended to constructing a confidence interval for a regression coefficient.
- Using a two-sided test, a hypothesized value for β_1 will be rejected at *5% significance level* if $\mid t^{act} \mid > critical\ value.1.96$.
- So and will be in the confidence set if $\mid t^{act} \mid \leq critical\ value.1.96$.
- Thus the 95% confidence interval for β_1 are within ± 1.96 standard errors of $\hat{\beta}_1$

$$\hat{\beta}_1 \pm 1.96 \cdot SE\left(\hat{\beta}_1\right)$$

Confidence interval for a regression coefficient β_1

- Method for constructing a confidence interval for a population mean can be easily extended to constructing a confidence interval for a regression coefficient.
- Using a two-sided test, a hypothesized value for β_1 will be rejected at *5% significance level* if $| t^{act} | > critical\ value.1.96$.
- So and will be in the confidence set if $| t^{act} | \leq critical\ value.1.96$.
- Thus the 95% confidence interval for β_1 are within ± 1.96 standard errors of $\hat{\beta}_1$

$$\hat{\beta}_1 \pm 1.96 \cdot SE\left(\hat{\beta}_1\right)$$

Confidence interval for a regression coefficient β_1

- Method for constructing a confidence interval for a population mean can be easily extended to constructing a confidence interval for a regression coefficient.
- Using a two-sided test, a hypothesized value for β_1 will be rejected at *5% significance level* if $\mid t^{act} \mid > critical\, value.1.96$.
- So and will be in the confidence set if $\mid t^{act} \mid \leq critical\, value.1.96$.
- Thus the 95% confidence interval for β_1 are within ± 1.96 standard errors of $\hat{\beta}_1$

$$\hat{\beta}_1 \pm 1.96 \cdot SE\left(\hat{\beta}_1\right)$$

Confidence interval for $\beta_{ClassSize}$

- Thus the 95% confidence interval for β_1 are within ± 1.96 standard errors of $\hat{\beta}_1$

$$\hat{\beta}_1 \pm 1.96 \cdot SE\left(\hat{\beta}_1\right) = -2.28 \pm (1.96 \times 0.52) = [-3.3, -1.26]$$

```
. regress test_score class_size, robust

Linear regression                               Number of obs     =        420
                                                  F(1, 418)         =      19.26
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.0512
                                                  Root MSE          =     18.581
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-2.279808	.5194892	-4.39	0.000	-3.300945	-1.258671
_cons	698.933	10.36436	67.44	0.000	678.5602	719.3057

Properties of Y as estimator of μ_Y

- Recall we discussed the properties of $\bar{Y}$ in *Chapter 2.* It is
  - an unbiased estimator of μ_Y
  - a consistent estimator of μ_Y
  - has an approximate normal sampling distribution for large n
  - the Best Linear Unbiased Estimator(BLUE): it is the most efficient estimator of μ_Y among all unbiased estimators.

Properties of Y as estimator of μ_Y

- Recall we discussed the properties of $\bar{Y}$ in *Chapter 2.* It is
  - an unbiased estimator of μ_Y
  - a consistent estimator of μ_Y
  - has an approximate normal sampling distribution for large n
  - the Best Linear Unbiased Estimator(BLUE): it is the most efficient estimator of μ_Y among all unbiased estimators.

Properties of Y as estimator of μ_Y

- Recall we discussed the properties of $\bar{Y}$ in *Chapter 2.* It is
  - an unbiased estimator of μ_Y
  - a consistent estimator of μ_Y
  - has an approximate normal sampling distribution for large n
  - the Best Linear Unbiased Estimator(BLUE): it is the most efficient estimator of μ_Y among all unbiased estimators.

Properties of Y as estimator of μ_Y

- Recall we discussed the properties of $\bar{Y}$ in *Chapter 2.* It is
  - an unbiased estimator of μ_Y
  - a consistent estimator of μ_Y
  - has an approximate normal sampling distribution for large n
  - the Best Linear Unbiased Estimator(BLUE): it is the most efficient estimator of μ_Y among all unbiased estimators.

Properties of Y as estimator of μ_Y

- Recall we discussed the properties of $\bar{Y}$ in *Chapter 2.* It is
  - an unbiased estimator of μ_Y
  - a consistent estimator of μ_Y
  - has an approximate normal sampling distribution for large n
  - the Best Linear Unbiased Estimator(BLUE): it is the most efficient estimator of μ_Y among all unbiased estimators.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - **Assumption 4**: The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - Assumption 4: The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - **Assumption 4:** The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - **Assumption 4**: The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - **Assumption 4**: The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - **Assumption 4**: The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Gauss-Markov theorem(SW: 5.5)

- Three Basic Assumption PLUS homoskedastic assumption, thus
  - Assumption 1
  - Assumption 2
  - Assumption 3
- we add a fourth OLS assumption:
  - **Assumption 4**: The error terms are homoskedastic

$$Var(u_i \mid X_i) = \sigma_u^2$$

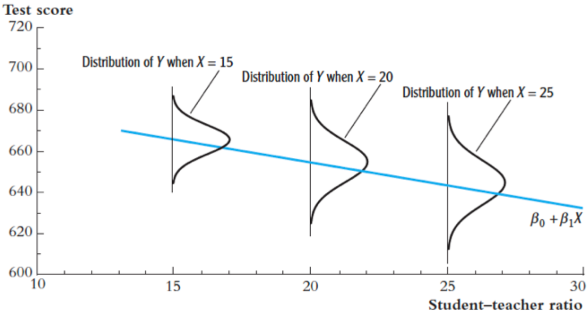
- Then $\hat{\beta}^{OLS}$ is the Best Linear Unbiased Estimator (BLUE): it is the most efficient estimator of β_1 among all conditional unbiased estimators that are a linear function of $Y_1, Y_2, ..., Y_n$.

Heteroskedasticity & homoskedasticity

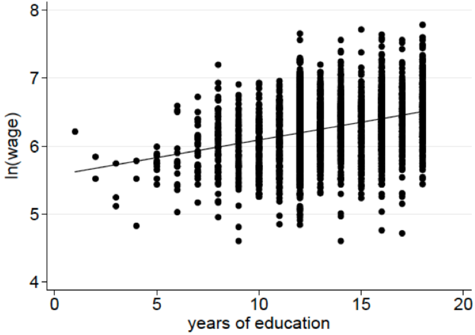
- The error term u_i is **homoskedastic** if the variance of the conditional distribution of u_i given X_i is constant for $i = 1, \dots n$, in particular does not depend on X_i. Otherwise, the error term is **heteroskedastic**.

FIGURE 5.2 An Example of Heteroskedasticity

Like Figure 4.4, this shows the conditional distribution of test scores for three different class sizes. Unlike Figure 4.4, these distributions become more spread out (have a larger variance) for larger class sizes. Because the variance of the distribution of u given X, var($u|X$), depends on X, u is heteroskedastic.

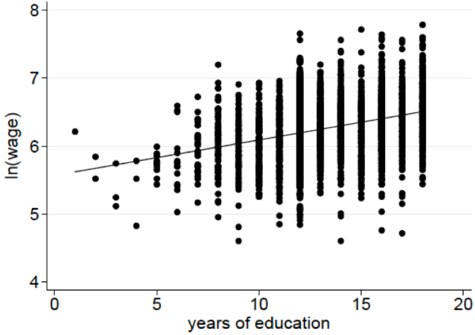


An Example: the returns to schooling



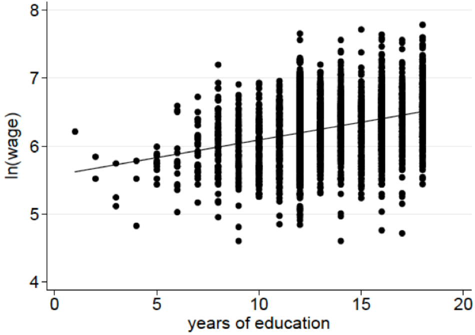
- The spread of the dots around the line is clearly increasing with years of education X_i.
- Variation in (log) wages is higher at higher levels of education.
- This implies that $Var(u_i \mid X_i) \neq \sigma_u^2$.

An Example: the returns to schooling



- The spread of the dots around the line is clearly increasing with years of education X_i.
- Variation in (log) wages is higher at higher levels of education.
- This implies that $Var(u_i \mid X_i) \neq \sigma_u^2$.

An Example: the returns to schooling



- The spread of the dots around the line is clearly increasing with years of education X_i.
- Variation in (log) wages is higher at higher levels of education.
- This implies that $Var(u_i \mid X_i) \neq \sigma_u^2$.

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic we should use the following *heteroskedasticity robust standard errors*

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

- If we assume that the error terms are *homoskedastic* the standard errors of the OLS estimators simplify to

$$SE\left(\hat{\beta}_1\right) = \sqrt{\frac{s^2_{\hat{u}}}{\sum(X_i - \bar{X})^2}}$$

- In many applications homoskedasticity is not a plausible assumption.If the error terms are heteroskedastic, then you use the homoskedastic assumption to compute the S.E. of $\hat{\beta}_1$.
  - The standard errors are wrong (often too small)
  - The t-statistic does NOT have a $N(0,1)$distribution (also not in large samples).

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic we should use the following *heteroskedasticity robust standard errors*

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

- If we assume that the error terms are *homoskedastic* the standard errors of the OLS estimators simplify to

$$SE\left(\hat{\beta}_1\right) = \sqrt{\frac{s^2_{\hat{u}}}{\sum(X_i - \bar{X})^2}}$$

- In many applications homoskedasticity is not a plausible assumption.If the error terms are heteroskedastic, then you use the homoskedastic assumption to compute the S.E. of $\hat{\beta}_1$.
  - The standard errors are wrong (often too small).
  - The t-statistic does NOT have a $N(0,1)$distribution (also not in large samples).

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic we should use the following *heteroskedasticity robust standard errors*

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

- If we assume that the error terms are *homoskedastic* the standard errors of the OLS estimators simplify to

$$SE\left(\hat{\beta}_1\right) = \sqrt{\frac{s^2_{\hat{u}}}{\sum(X_i - \bar{X})^2}}$$

- In many applications homoskedasticity is not a plausible assumption.If the error terms are heteroskedastic, then you use the homoskedastic assumption to compute the S.E. of $\hat{\beta}_1$.
  - The standard errors are wrong (often too small)
  - The t-statistic does NOT have a $N(0,1)$distribution (also not in large samples).

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic we should use the following *heteroskedasticity robust standard errors*

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

- If we assume that the error terms are *homoskedastic* the standard errors of the OLS estimators simplify to

$$SE\left(\hat{\beta}_1\right) = \sqrt{\frac{s^2_{\hat{u}}}{\sum(X_i - \bar{X})^2}}$$

- In many applications homoskedasticity is not a plausible assumption.If the error terms are heteroskedastic, then you use the homoskedastic assumption to compute the S.E. of $\hat{\beta}_1$.
  - The standard errors are wrong (often too small)
  - The t-statistic does NOT have a $N(0,1)$distribution (also not in large samples).

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic we should use the following *heteroskedasticity robust standard errors*

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

- If we assume that the error terms are *homoskedastic* the standard errors of the OLS estimators simplify to

$$SE\left(\hat{\beta}_1\right) = \sqrt{\frac{s^2_{\hat{u}}}{\sum(X_i - \bar{X})^2}}$$

- In many applications homoskedasticity is not a plausible assumption.If the error terms are heteroskedastic, then you use the homoskedastic assumption to compute the S.E. of $\hat{\beta}_1$.
    - The standard errors are wrong (often too small)
    - The t-statistic does NOT have a $N(0,1)$distribution (also not in large samples).

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic we should use the following *heteroskedasticity robust standard errors*

$$SE\left(\hat{\beta}_1\right) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\frac{1}{n} \times \frac{\frac{1}{n-2}\sum(X_i - \bar{X})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^2}}$$

- If we assume that the error terms are *homoskedastic* the standard errors of the OLS estimators simplify to

$$SE\left(\hat{\beta}_1\right) = \sqrt{\frac{s^2_{\hat{u}}}{\sum(X_i - \bar{X})^2}}$$

- In many applications homoskedasticity is not a plausible assumption.If the error terms are heteroskedastic, then you use the homoskedastic assumption to compute the S.E. of $\hat{\beta}_1$.
  - The standard errors are wrong (often too small)
  - The t-statistic does NOT have a $N(0,1)$distribution (also not in large samples).

Heteroskedasticity & homoskedasticity

- Since homoskedasticity is a **special case** of heteroskedasticity, these heteroskedasticity robust formulas are also valid if the error terms are homoskedastic.
- Hypothesis tests and confidence intervals based on above SE' s are valid both in case of homoskedasticity and heteroskedasticity.
- In reality, since in many applications homoskedasticity is not a plausible assumption It is best to use heteroskedasticity robust standard errors. (we lose nothing)
- In ***Stata*** the default option of regression is to assume homoskedasticity, to obtain heteroskedasticity robust standard errors use the option "**robust**" :

$$Regress\ y\ x\ ,\ robust$$

Heteroskedasticity & homoskedasticity

- Since homoskedasticity is a **special case** of heteroskedasticity, these heteroskedasticity robust formulas are also valid if the error terms are homoskedastic.
- Hypothesis tests and confidence intervals based on above SE's are valid both in case of homoskedasticity and heteroskedasticity.
- In reality, since in many applications homoskedasticity is not a plausible assumption It is best to use heteroskedasticity robust standard errors. (we lose nothing)
- In ***Stata*** the default option of regression is to assume homoskedasticity, to obtain heteroskedasticity robust standard errors use the option "**robust**":

$$Regress\ y\ x\ ,\ robust$$

Heteroskedasticity & homoskedasticity

- Since homoskedasticity is a **special case** of heteroskedasticity, these heteroskedasticity robust formulas are also valid if the error terms are homoskedastic.
- Hypothesis tests and confidence intervals based on above SE' s are valid both in case of homoskedasticity and heteroskedasticity.
- In reality, since in many applications homoskedasticity is not a plausible assumption It is best to use heteroskedasticity robust standard errors. (we lose nothing)
- In ***Stata*** the default option of regression is to assume homoskedasticity, to obtain heteroskedasticity robust standard errors use the option "**robust**":

$$Regress\ y\ x\ ,\ robust$$

Heteroskedasticity & homoskedasticity

- Since homoskedasticity is a **special case** of heteroskedasticity, these heteroskedasticity robust formulas are also valid if the error terms are homoskedastic.
- Hypothesis tests and confidence intervals based on above SE' s are valid both in case of homoskedasticity and heteroskedasticity.
- In reality, since in many applications homoskedasticity is not a plausible assumption It is best to use heteroskedasticity robust standard errors. (we lose nothing)
- In ***Stata*** the default option of regression is to assume homoskedasticity, to obtain heteroskedasticity robust standard errors use the option "**robust**" :

$$Regress\ y\ x\ ,\ \ robust$$

Heteroskedasticity & homoskedasticity

```
. regress test_score class_size

      Source |       SS           df       MS      Number of obs   =       420
-------------+----------------------------------   F(1, 418)       =     22.58
       Model |  7794.11004         1  7794.11004   Prob > F        =    0.0000
    Residual |  144315.484       418  345.252353   R-squared       =    0.0512
-------------+----------------------------------   Adj R-squared   =    0.0490
       Total |  152109.594       419  363.030056   Root MSE        =    18.581

------------------------------------------------------------------------------
  test_score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  class_size |  -2.279808   .4798256    -4.75   0.000     -3.22298   -1.336637
       _cons |    698.933   9.467491    73.82   0.000     680.3231    717.5428
------------------------------------------------------------------------------

. regress test_score class_size, robust

Linear regression                               Number of obs     =        420
                                                F(1, 418)         =      19.26
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0512
                                                Root MSE          =     18.581

------------------------------------------------------------------------------
             |               Robust
  test_score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  class_size |  -2.279808   .5194892    -4.39   0.000    -3.300945   -1.258671
       _cons |    698.933   10.36436    67.44   0.000     678.5602    719.3057
------------------------------------------------------------------------------
```

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

Heteroskedasticity & homoskedasticity

- If the error terms are heteroskedastic
  - The fourth OLS assumption is violated
  - The Gauss-Markov conditions do not hold
  - The OLS estimator is not BLUE (not efficient)
- But (given that the other OLS assumptions hold)
  - The OLS estimators are unbiased
  - The OLS estimators are consistent
  - The OLS estimators are normally distributed in large samples

OLS with Multiple Regressors: Hypotheses tests

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Least Squares assumptions of the multiple regression

- Fourth Basic Assumption
  - Assumption 1 : $E[u_i \mid X_{1i}, X_{2i}..., X_{ki}] = 0$
  - Assumption 2 : i.i.d sample
  - Assumption 3 : Large outliers are unlikely.
  - Assumption 4: No perfect multicollinearity.
- the OLS estimators $\hat{\beta}_j$ for $j = 1, ..., k$ are approximately normally distributed in large samples. In addition

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)} \sim N(0,1)$$

- We can thus perform, hypothesis tests in same way as in regression model with only one regressor.

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $| t^{act} | > critical\ value$
    - or if $p - value < significance\ level$

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $|t^{act}| > critical\ value$
    - or if $p - value < significance\ level$

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Hypothesis test for single coefficient

- $H_0 : \beta_j = \beta_{j,0}$ $H_1 : \beta_1 \neq \beta_{j,0}$
  - Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS to obtain $\hat{\beta}_j$
  - Step2: Compute the **standard error** of $\hat{\beta}_j$ (requires matrix algebra)
  - Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_j - \beta_{j,0}}{SE\left(\hat{\beta}_j\right)}$$

  - Step4: Reject the null hypothesis if
    - $\mid t^{act} \mid > critical\ value$
    - or if $p - value < significance\ level$

Case: Class Size and test scores

```
. regress test_score class_size el_pct, robust

Linear regression                               Number of obs     =        420
                                                  F(2, 417)         =     223.82
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.4264
                                                  Root MSE          =     14.464
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-1.101296	.4328472	-2.54	0.011	-1.95213	-.2504616
el_pct	-.6497768	.0310318	-20.94	0.000	-.710775	-.5887786
_cons	686.0322	8.728224	78.60	0.000	668.8754	703.189

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0$ $H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\ value 1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \ H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\ value.1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \; H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\ value.1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \ H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\ value.1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \ H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $| \ t^{act} \ | = | -2.54 | > critical \ value 1.96$
  - $p - value = 0.011 < significance \ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \ H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\ value.1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \; H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\,value.1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Case: Class Size and test scores

- Does changing class size, *while holding the percentage of English learners constant*, have a statistically significant effect on test scores? (using a 5% significance level)
- $H_0 : \beta_{ClassSize} = 0 \ H_1 : \beta_{ClassSize} \neq 0$
- Step1: Estimate $\hat{\beta}_1 = -1.10$
- Step2: Compute the standard error: $SE(\hat{\beta}_1) = 0.43$
- Step3: Compute the t-statistic

$$t^{act} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} = \frac{-1.10 - 0}{0.43} = -2.54$$

- Step4: Reject the null hypothesis if
  - $\mid t^{act} \mid = \mid -2.54 \mid > critical\ value.1.96$
  - $p - value = 0.011 < significance\ level = 0.05$

Testing 1 hypothesis on 2 or more coefficients

- Suppose we want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
- $H_0: \beta_{meal\,pct} = 0 \; \& \beta_{calw\,pct} = 0$
  $H_1: \beta_{meal\,pct} \neq 0 \; and/or \; \beta_{calw\,pct} \neq 0$
- If either $t_{meal\,pct}$ or $t_{calw\,pct}$ exceeds 1.96, we should reject?
- We assume that $t_{meal\,pct}$ and $t_{calw\,pct}$ are *uncorrelated*:

$$\begin{aligned} Pr(t_{meal\,pct} > 1.96 \; and/or \; t_{calw\,pct} > 1.96) &= 1 - Pr(t_{meal\,pct} > 1.96 \\ &= 1 - Pr(t_{meal\,pct} > 1.96) \\ &= 1 - 0.95 \times 0.95 \\ &= 0.0975 > 0.05 \end{aligned}$$

- if $t_{meal\,pct}$ and $t_{calw\,pct}$ are *correlated, then it is more complicated.*

Testing 1 hypothesis on 2 or more coefficients

- Suppose we want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
- $H_0 : \beta_{meal\,pct} = 0 \;\& \beta_{calw\,pct} = 0$
  $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \; \beta_{calw\,pct} \neq 0$
- If either $t_{meal\,pct}$ or $t_{calw\,pct}$ exceeds 1.96, we should reject?
- We assume that $t_{meal\,pct}$ and $t_{calw\,pct}$ are *uncorrelated*:

$$\begin{aligned} Pr(t_{meal\,pct} > 1.96 \; and/or \; t_{calw\,pct} > 1.96) &= 1 - Pr(t_{meal\,pct} > 1.96 \\ &= 1 - Pr(t_{meal\,pct} > 1.96) \\ &= 1 - 0.95 \times 0.95 \\ &= 0.0975 > 0.05 \end{aligned}$$

- if $t_{meal\,pct}$ and $t_{calw\,pct}$ are *correlated, then it is more complicated.*

Testing 1 hypothesis on 2 or more coefficients

- Suppose we want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
- $H_0 : \beta_{meal\,pct} = 0\ \&\beta_{calw\,pct} = 0$
  $H_1 : \beta_{meal\,pct} \neq 0\ and/or\ \beta_{calw\,pct} \neq 0$
- If either $t_{meal\,pct}$ or $t_{calw\,pct}$ exceeds 1.96, we should reject?
- We assume that $t_{meal\,pct}$ and $t_{calw\,pct}$ are *uncorrelated*:

$$\begin{aligned} Pr(t_{meal\,pct} > 1.96\ and/or\ t_{calw\,pct} > 1.96) &= 1 - Pr(t_{meal\,pct} > 1.96 \\ &= 1 - Pr(t_{meal\,pct} > 1.96) \\ &= 1 - 0.95 \times 0.95 \\ &= 0.0975 > 0.05 \end{aligned}$$

- if $t_{meal\,pct}$ and $t_{calw\,pct}$ are *correlated, then it is more complicated.*

Testing 1 hypothesis on 2 or more coefficients

- Suppose we want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
- $H_0 : \beta_{meal\,pct} = 0 \,\&\beta_{calw\,pct} = 0$
  $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \, \beta_{calw\,pct} \neq 0$
- If either $t_{meal\,pct}$ or $t_{calw\,pct}$ exceeds 1.96, we should reject?
- We assume that $t_{meal\,pct}$ and $t_{calw\,pct}$ are *uncorrelated*:

$$
\begin{aligned}
Pr(t_{meal\,pct} > 1.96 \; and/or \, t_{calw\,pct} > 1.96) &= 1 - Pr(t_{meal\,pct} > 1.96 \\
&= 1 - Pr(t_{meal\,pct} > 1.96) \\
&= 1 - 0.95 \times 0.95 \\
&= 0.0975 > 0.05
\end{aligned}
$$

- if $t_{meal\,pct}$ and $t_{calw\,pct}$ are *correlated, then it is more complicated.*

Testing 1 hypothesis on 2 or more coefficients

- Suppose we want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
- $H_0 : \beta_{meal\,pct} = 0\ \&\beta_{calw\,pct} = 0$
  $H_1 : \beta_{meal\,pct} \neq 0\ and/or\,\beta_{calw\,pct} \neq 0$
- If either $t_{meal\,pct}$ or $t_{calw\,pct}$ exceeds 1.96, we should reject?
- We assume that $t_{meal\,pct}$ and $t_{calw\,pct}$ are *uncorrelated*:

$$
\begin{aligned}
Pr(t_{meal\,pct} > 1.96\ and/or\,t_{calw\,pct} > 1.96) &= 1 - Pr(t_{meal\,pct} > 1.96 \\
&= 1 - Pr(t_{meal\,pct} > 1.96) \\
&= 1 - 0.95 \times 0.95 \\
&= 0.0975 > 0.05
\end{aligned}
$$

- if $t_{meal\,pct}$ and $t_{calw\,pct}$ are *correlated, then it is more complicated.*

Heteroskedasticity & homoskedasticity

- If we want to test joint hypotheses that involves multiple coefficients we need to use an **F-test** based on the **F-statistic**
- F-Statistic with $q = 2$: when testing the following hypothesis

$$H_0 : \beta_1 = 0 \;\&\; \beta_2 = 0 \quad H_1 : \beta_1 \neq 0 \; and/or \; \beta_2 \neq 0$$

- the F-statistic combines the two t-statistics as follows

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1 t_2} t_1 t_2}{1 - \hat{\rho}_{t_1 t_2}^2}\right)$$

where $\hat{\rho}_{t_1 t_2}$ is an estimator of the correlation between the two t-statistics.

Heteroskedasticity & homoskedasticity

- If we want to test joint hypotheses that involves multiple coefficients we need to use an **F-test** based on the **F-statistic**
- F-Statistic with $q = 2$: when testing the following hypothesis

$$H_0 : \beta_1 = 0 \;\&\; \beta_2 = 0 \quad H_1 : \beta_1 \neq 0 \; and/or \; \beta_2 \neq 0$$

- the F-statistic combines the two t-statistics as follows

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1 t_2} t_1 t_2}{1 - \hat{\rho}_{t_1 t_2}^2}\right)$$

where $\hat{\rho}_{t_1 t_2}$ is an estimator of the correlation between the two t-statistics.

Heteroskedasticity & homoskedasticity

- If we want to test joint hypotheses that involves multiple coefficients we need to use an **F-test** based on the **F-statistic**
- F-Statistic with $q = 2$: when testing the following hypothesis

$$H_0 : \beta_1 = 0 \;\&\; \beta_2 = 0 \quad H_1 : \beta_1 \neq 0 \; and/or \; \beta_2 \neq 0$$

- the F-statistic combines the two t-statistics as follows

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1 t_2} t_1 t_2}{1 - \hat{\rho}^2_{t_1 t_2}}\right)$$

where $\hat{\rho}_{t_1 t_2}$is an estimator of the correlation between the two t-statistics.

Case: Class Size and test scores

- We want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
  - $H_0: \beta_{meal\,pct} = 0 \,\&\beta_{calw\,pct} = 0$
    $H_1: \beta_{meal\,pct} \neq 0 \; and/or \, \beta_{calw\,pct} \neq 0$
- The null hypothesis consists of two restrictions q = 2
- It can be shown that the F-statistic with two restrictions has an approximate $F_{2,\infty}$ distribution *in large samples*

$$F = 290.27$$

- Table 4 (S&W page 795) shows that the critical value at a 5% significance level equals 3.00
- This implies that we reject H_0 at a 5% significance level because $290.27 > 3$

Case: Class Size and test scores

- We want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
  - $H_0 : \beta_{meal\,pct} = 0 \,\& \beta_{calw\,pct} = 0$
    $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \, \beta_{calw\,pct} \neq 0$
- The null hypothesis consists of two restrictions q = 2
- It can be shown that the F-statistic with two restrictions has an approximate $F_{2,\infty}$ distribution *in large samples*

$$F = 290.27$$

- Table 4 (S&W page 795) shows that the critical value at a 5% significance level equals 3.00
- This implies that we reject H_0 at a 5% significance level because $290.27 > 3$

Case: Class Size and test scores

- We want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
  - $H_0 : \beta_{meal\,pct} = 0 \, \& \beta_{calw\,pct} = 0$
    $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \; \beta_{calw\,pct} \neq 0$
- The null hypothesis consists of two restrictions q = 2
- It can be shown that the F-statistic with two restrictions has an approximate $F_{2,\infty}$ distribution *in large samples*

$$F = 290.27$$

- Table 4 (S&W page 795) shows that the critical value at a 5% significance level equals 3.00
- This implies that we reject H_0 at a 5% significance level because $290.27 > 3$

Case: Class Size and test scores

- We want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
  - $H_0 : \beta_{meal\,pct} = 0 \,\& \beta_{calw\,pct} = 0$
    $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \; \beta_{calw\,pct} \neq 0$
- The null hypothesis consists of two restrictions q $= 2$
- It can be shown that the F-statistic with two restrictions has an approximate $F_{2,\infty}$ distribution *in large samples*

$$F = 290.27$$

- Table 4 (S&W page 795) shows that the critical value at a 5% significance level equals 3.00
- This implies that we reject H_0 at a 5% significance level because $290.27 > 3$

Case: Class Size and test scores

- We want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
  - $H_0 : \beta_{meal\,pct} = 0 \,\&\beta_{calw\,pct} = 0$
    $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \; \beta_{calw\,pct} \neq 0$
- The null hypothesis consists of two restrictions q $= 2$
- It can be shown that the F-statistic with two restrictions has an approximate $F_{2,\infty}$ distribution *in large samples*

$$F = 290.27$$

- Table 4 (S&W page 795) shows that the critical value at a 5% significance level equals 3.00
- This implies that we reject H_0 at a 5% significance level because $290.27 > 3$

Case: Class Size and test scores

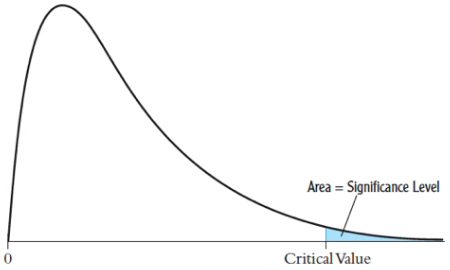
- We want to test hypothesis that both the coefficient on % eligible for a free lunch and the coefficient on % eligible for calworks are zero?
  - $H_0 : \beta_{meal\,pct} = 0 \,\& \beta_{calw\,pct} = 0$
    $H_1 : \beta_{meal\,pct} \neq 0 \; and/or \; \beta_{calw\,pct} \neq 0$
- The null hypothesis consists of two restrictions q $= 2$
- It can be shown that the F-statistic with two restrictions has an approximate $F_{2,\infty}$ distribution *in large samples*

$$F = 290.27$$

- Table 4 (S&W page 795) shows that the critical value at a 5% significance level equals 3.00
- This implies that we reject H_0 at a 5% significance level because $290.27 > 3$

F-Test

TABLE 4 Critical Values for the $F_{m,\infty}$ Distribution



	Significance Level		
Degrees of Freedom	**10%**	**5%**	**1%**
1	2.71	3.84	6.63
2	2.30	3.00	4.61
3	2.08	2.60	3.78

General procedure for testing joint hypothesis with q restrictions

- $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of q restrictions.
- H_1 :at least one of q restrictions under H_0 does not hold.
- Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS
- Step2: Compute the F-statistic
- Step3 : Reject the null hypothesis if $F - Statistic > F^{act}_{q,\infty}$ or $p - value = Pr[F_{q,\infty} > F^{act}]$

General procedure for testing joint hypothesis with q restrictions

- $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of q restrictions.
- H_1 :at least one of q restrictions under H_0 does not hold.
- Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS
- Step2: Compute the F-statistic
- Step3 : Reject the null hypothesis if $F-Statistic > F^{act}_{q,\infty}$ or $p-value = Pr[F_{q,\infty} > F^{act}]$

General procedure for testing joint hypothesis with q restrictions

- $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of q restrictions.
- H_1 :at least one of q restrictions under H_0 does not hold.
- Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS
- Step2: Compute the F-statistic
- Step3 : Reject the null hypothesis if $F - Statistic > F^{act}_{q,\infty}$ or $p - value = Pr[F_{q,\infty} > F^{act}]$

General procedure for testing joint hypothesis with q restrictions

- $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of q restrictions.
- H_1 :at least one of q restrictions under H_0 does not hold.
- Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS
- Step2: Compute the F-statistic
- Step3 : Reject the null hypothesis if $F - Statistic > F^{act}_{q,\infty}$ or $p - value = Pr[F_{q,\infty} > F^{act}]$

General procedure for testing joint hypothesis with q restrictions

- $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of q restrictions.
- H_1 :at least one of q restrictions under H_0 does not hold.
- Step1: Estimate $Y_i = \beta_0 + \beta_1 X_{1i} + ... + \beta_j X_{ji} + ... + \beta_k X_{ki} + u_i$ by OLS
- Step2: Compute the F-statistic
- Step3 : Reject the null hypothesis if $F - Statistic > F^{act}_{q,\infty}$ or $p - value = Pr[F_{q,\infty} > F^{act}]$

Case: Class Size and test scores: q=3 restrictions

```
1 . regress test_score class_size el_pct meal_pct calw_pct, robust

  Linear regression                              Number of obs     =        420
                                                   F(4, 415)         =     361.68
                                                   Prob > F          =     0.0000
                                                   R-squared         =     0.7749
                                                   Root MSE          =     9.0843

  ------------------------------------------------------------------------------
               |               Robust
    test_score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
  -------------+----------------------------------------------------------------
    class_size |  -1.014353   .2688613    -3.77   0.000    -1.542853   -.4858534
        el_pct |  -.1298219   .0362579    -3.58   0.000     -.201094   -.0585498
      meal_pct |  -.5286191   .0381167   -13.87   0.000    -.6035449   -.4536932
      calw_pct |  -.0478537   .0586541    -0.82   0.415    -.1631498    .0674424
         _cons |   700.3918   5.537418   126.48   0.000      689.507    711.2767
  ------------------------------------------------------------------------------

2 . test el_pct meal_pct calw_pct

   ( 1)  el_pct = 0
   ( 2)  meal_pct = 0
   ( 3)  calw_pct = 0

         F(  3,   415) =  481.06
              Prob > F =    0.0000
```

-

Case: Class Size and test scores: q=3 restrictions

- $H_0 : \beta_{el\,pct} = \beta_{meal\,pct} = \beta_{calw\,pct} = 0$
- H_1 :at least one of q restrictions under H_0 does not hold.
  1. Step1: Estimate by OLS
  2. Step2: $F - Statistic = 481.06$
  3. Step3: We reject the null hypothesis at a 5% significance level because $F - Statistic > F_{3,\infty} = 2.6$

Case: Class Size and test scores: q=3 restrictions

- $H_0 : \beta_{el\,pct} = \beta_{meal\,pct} = \beta_{calw\,pct} = 0$
- H_1 :at least one of q restrictions under H_0 does not hold.
  1. Step1: Estimate by OLS
  2. Step2: $F - Statistic = 481.06$
  3. Step3: We reject the null hypothesis at a 5% significance level because $F - Statistic > F_{3,\infty} = 2.6$

Case: Class Size and test scores: q=3 restrictions

- $H_0 : \beta_{el\,pct} = \beta_{meal\,pct} = \beta_{calw\,pct} = 0$
- H_1 :at least one of q restrictions under H_0 does not hold.
    1. Step1: Estimate by OLS
    2. Step2: $F - Statistic = 481.06$
    3. Step3: We reject the null hypothesis at a 5% significance level because $F - Statistic > F_{3,\infty} = 2.6$

Case: Class Size and test scores: q=3 restrictions

- $H_0 : \beta_{el\,pct} = \beta_{meal\,pct} = \beta_{calw\,pct} = 0$
- H_1 :at least one of q restrictions under H_0 does not hold.
  1. Step1: Estimate by OLS
  2. Step2: $F - Statistic = 481.06$
  3. Step3: We reject the null hypothesis at a 5% significance level because $F - Statistic > F_{3,\infty} = 2.6$

Case: Class Size and test scores: q=3 restrictions

- $H_0 : \beta_{el\,pct} = \beta_{meal\,pct} = \beta_{calw\,pct} = 0$
- H_1 :at least one of q restrictions under H_0 does not hold.
  1. Step1: Estimate by OLS
  2. Step2: $F - Statistic = 481.06$
  3. Step3: We reject the null hypothesis at a 5% significance level because $F - Statistic > F_{3,\infty} = 2.6$

The "overall" regression F-statistic

- The "overall" F-statistic test the joint hypothesis that all the k slope coefficients are zero
  - $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of $q = k$ restrictions.
  - H_1 : at least one of $q = k$ restrictions under H_0 does not hold.

```
. regress test_score class_size el_pct meal_pct calw_pct, robust

Linear regression                               Number of obs     =        420
                                                  F(4, 415)         =     361.68
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.7749
                                                  Root MSE          =     9.0843
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-1.014353	.2688613	-3.77	0.000	-1.542853	-.4858534
el_pct	-.1298219	.0362579	-3.58	0.000	-.201094	-.0585498
meal_pct	-.5286191	.0381167	-13.87	0.000	-.6035449	-.4536932
calw_pct	-.0478537	.0586541	-0.82	0.415	-.1631498	.0674424
_cons	700.3918	5.537418	126.48	0.000	689.507	711.2767

- The overall $F - Statistics = 361.68$

The "overall" regression F-statistic

- The "overall" F-statistic test the joint hypothesis that all the k slope coefficients are zero
  - $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of $q = k$ restrictions.
  - H_1 : at least one of $q = k$ restrictions under H_0 does not hold.

```
. regress test_score class_size el_pct meal_pct calw_pct, robust

Linear regression                               Number of obs     =        420
                                                  F(4, 415)         =     361.68
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.7749
                                                  Root MSE          =     9.0843
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-1.014353	.2688613	-3.77	0.000	-1.542853	-.4858534
el_pct	-.1298219	.0362579	-3.58	0.000	-.201094	-.0585498
meal_pct	-.5286191	.0381167	-13.87	0.000	-.6035449	-.4536932
calw_pct	-.0478537	.0586541	-0.82	0.415	-.1631498	.0674424
_cons	700.3918	5.537418	126.48	0.000	689.507	711.2767

- The overall $F - Statistics = 361.68$

The "overall" regression F-statistic

- The "overall" F-statistic test the joint hypothesis that all the k slope coefficients are zero
  - $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of $q = k$ restrictions.
  - H_1 : at least one of $q = k$ restrictions under H_0 does not hold.

```
. regress test_score class_size el_pct meal_pct calw_pct, robust

Linear regression                               Number of obs     =        420
                                                F(4, 415)         =     361.68
                                                Prob > F          =     0.0000
                                                R-squared         =     0.7749
                                                Root MSE          =     9.0843
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-1.014353	.2688613	-3.77	0.000	-1.542853	-.4858534
el_pct	-.1298219	.0362579	-3.58	0.000	-.201094	-.0585498
meal_pct	-.5286191	.0381167	-13.87	0.000	-.6035449	-.4536932
calw_pct	-.0478537	.0586541	-0.82	0.415	-.1631498	.0674424
_cons	700.3918	5.537418	126.48	0.000	689.507	711.2767

- The overall $F - Statistics = 361.68$

The "overall" regression F-statistic

- The "overall" F-statistic test the joint hypothesis that all the k slope coefficients are zero
  - $H_0 : \beta_j = \beta_{j,0}, ..., \beta_m = \beta_{m,0}$ for a total of $q = k$ restrictions.
  - H_1 : at least one of $q = k$ restrictions under H_0 does not hold.

```
. regress test_score class_size el_pct meal_pct calw_pct, robust

Linear regression                                Number of obs     =        420
                                                   F(4, 415)         =     361.68
                                                   Prob > F          =     0.0000
                                                   R-squared         =     0.7749
                                                   Root MSE          =     9.0843
```

test_score	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-1.014353	.2688613	-3.77	0.000	-1.542853	-.4858534
el_pct	-.1298219	.0362579	-3.58	0.000	-.201094	-.0585498
meal_pct	-.5286191	.0381167	-13.87	0.000	-.6035449	-.4536932
calw_pct	-.0478537	.0586541	-0.82	0.415	-.1631498	.0674424
_cons	700.3918	5.537418	126.48	0.000	689.507	711.2767

- The overall $F - Statistics = 361.68$

The "Star War" and Regression Table

Dependent variable: average test score in the district.

Regressor	(1)	(2)	(3)	(4)	(5)
Student–teacher ratio (X_1)	−2.28** (0.52)	−1.10* (0.43)	−1.00** (0.27)	−1.31* (0.34)	−1.01* (0.27)
Percent English learners (X_2)		−0.650** (0.031)	−0.122** (0.033)	−0.488** (0.030)	−0.130** (0.036)
Percent eligible for subsidized lunch (X_3)			−0.547* (0.024)		−0.529* (0.038)
Percent on public income assistance (X_4)				−0.790** (0.068)	0.048 (0.059)
Intercept	698.9** (10.4)	686.0** (8.7)	700.2** (5.6)	698.0** (6.9)	700.4** (5.5)
Summary Statistics					
SER	18.58	14.46	9.08	11.65	9.08
$\overline{R}^2$	0.049	0.424	0.773	0.626	0.773
n	420	420	420	420	420

These regressions were estimated using the data on K–8 school districts in California, described in Appendix (4.1). Heteroskedasticity-robust standard errors are given in parentheses under coefficients. The individual coefficient is statistically significant at the *5% level or **1% significance level using a two-sided test.