ECO 344 - Applied Econometrics I
Chapter 4

Muhammad Salman Khalid

School of Economics & Social Sciences

September 25, 2026



Outline

1. Inference: Introduction
2. Sampling Distribution of OLS Estimators
3. Testing Hypotheses About a Single Parameter
4. Confidence Intervals
5. Testing Hypotheses about a Single Linear Combination of Parameters
6. Testing Hypotheses about a Multiple Linear Combination of Parameters

Inference: Introduction

Inference

- What is Inference?

Inference

- What is Inference?
- As we know, that the almost all of the time, we have the sample and not population. Therefore, all the statistics, we obtain are estimators and not parameters.

Inference

- What is Inference?
- As we know, that the almost all of the time, we have the sample and not population. Therefore, all the statistics, we obtain are estimators and not parameters.
- The estimators are different from parameters, even if they are unbiased. Recall;

$$E[\hat{\beta}] = \beta$$

Inference

- What is Inference?
- As we know, that the almost all of the time, we have the sample and not population. Therefore, all the statistics, we obtain are estimators and not parameters.
- The estimators are different from parameters, even if they are unbiased. Recall;

$$E[\hat{\beta}] = \beta$$

- So, the expected value of our estimator is equal to the population parameter and not the individual value of estimator we obtained from the sample.

Inference

- What is Inference?
- As we know, that the almost all of the time, we have the sample and not population. Therefore, all the statistics, we obtain are estimators and not parameters.
- The estimators are different from parameters, even if they are unbiased. Recall;

$$E[\hat{\beta}] = \beta$$

- So, the expected value of our estimator is equal to the population parameter and not the individual value of estimator we obtained from the sample.
- However, we can test some statements about our population parameters using the estimators of our sample. This is called Inference.

Inference

- Inference is a vast area to study and it can be a separate branch of statistics or econometrics.

Inference

- Inference is a vast area to study and it can be a separate branch of statistics or econometrics.
- Generally, the inference includes;

Inference

- Inference is a vast area to study and it can be a separate branch of statistics or econometrics.
- Generally, the inference includes;
  - Hypothesis Testing - Testing whether the sample evidence is consistent with a hypothesized value of the population parameter.

Inference

- Inference is a vast area to study and it can be a separate branch of statistics or econometrics.
- Generally, the inference includes;
  - Hypothesis Testing - Testing whether the sample evidence is consistent with a hypothesized value of the population parameter.
  - Confidence Interval - A range of plausible values for the population parameter, built by a procedure that covers the true value in a known share of repeated samples.

Inference

- Inference is a vast area to study and it can be a separate branch of statistics or econometrics.
- Generally, the inference includes;
  - Hypothesis Testing - Testing whether the sample evidence is consistent with a hypothesized value of the population parameter.
  - Confidence Interval - A range of plausible values for the population parameter, built by a procedure that covers the true value in a known share of repeated samples.
- Before we dive into inference, we need to know the sampling distribution of the OLS estimators.

Sampling Distribution of OLS Estimators

Sampling Distribution of OLS Estimators

- So far, we have the OLS estimators which are unbiased under M.L.R - I through M.L.R - IV. Under these four assumptions we know the average/expected value of OLS estimators.

Sampling Distribution of OLS Estimators

- So far, we have the OLS estimators which are unbiased under M.L.R - I through M.L.R - IV. Under these four assumptions we know the average/expected value of OLS estimators.
- Similarly, we also have the variance of OLS estimators and under M.L.R - I through M.L.R - V, the estimator of the error variance, $\hat{\sigma}^2$, is also unbiased. Under these five assumptions we have the standard error of the estimators.

Sampling Distribution of OLS Estimators

- So far, we have the OLS estimators which are unbiased under M.L.R - I through M.L.R - IV. Under these four assumptions we know the average/expected value of OLS estimators.
- Similarly, we also have the variance of OLS estimators and under M.L.R - I through M.L.R - V, the estimator of the error variance, $\hat{\sigma}^2$, is also unbiased. Under these five assumptions we have the standard error of the estimators.
- However, we have not made any assumption about the distribution of OLS estimators. For inference, we need the distribution.

Sampling Distribution of OLS Estimators

- So far, we have the OLS estimators which are unbiased under M.L.R - I through M.L.R - IV. Under these four assumptions we know the average/expected value of OLS estimators.
- Similarly, we also have the variance of OLS estimators and under M.L.R - I through M.L.R - V, the estimator of the error variance, $\hat{\sigma}^2$, is also unbiased. Under these five assumptions we have the standard error of the estimators.
- However, we have not made any assumption about the distribution of OLS estimators. For inference, we need the distribution.
- Therefore, we will make an assumption about the distribution of OLS estimators.

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

- M.L.R - VI assumption is a stricter assumption. It includes;

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

- M.L.R - VI assumption is a stricter assumption. It includes;
  1. M.L.R - IV - $E[\mu|x_1, x_2, x_3, ..., x_k] = E[\mu] = 0$

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

- M.L.R - VI assumption is a stricter assumption. It includes;
  1. M.L.R - IV - $E[\mu|x_1, x_2, x_3, ..., x_k] = E[\mu] = 0$
  2. M.L.R - V - $VAR[\mu|x_1, x_2, x_3, ..., x_k] = VAR[\mu] = \sigma^2$

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

- M.L.R - VI assumption is a stricter assumption. It includes;
  1. M.L.R - IV - $E[\mu|x_1, x_2, x_3, ..., x_k] = E[\mu] = 0$
  2. M.L.R - V - $VAR[\mu|x_1, x_2, x_3, ..., x_k] = VAR[\mu] = \sigma^2$
  3. The errors are normally distributed.

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

- M.L.R - VI assumption is a stricter assumption. It includes;
  1. M.L.R - IV - $E[\mu|x_1, x_2, x_3, ..., x_k] = E[\mu] = 0$
  2. M.L.R - V - $VAR[\mu|x_1, x_2, x_3, ..., x_k] = VAR[\mu] = \sigma^2$
  3. The errors are normally distributed.
- The assumption M.L.R - I through M.L.R - VI are known as **Classic Linear Model(CLM) Assumptions**

Normal Distribution of OLS Estimators

- The population error is independent of explanatory variables and is normally distributed with zero mean and variance of σ^2. - M.L.R - VI

$$\mu \sim Normal(0, \sigma^2)$$

- M.L.R - VI assumption is a stricter assumption. It includes;
  1. M.L.R - IV - $E[\mu|x_1, x_2, x_3, ..., x_k] = E[\mu] = 0$
  2. M.L.R - V - $VAR[\mu|x_1, x_2, x_3, ..., x_k] = VAR[\mu] = \sigma^2$
  3. The errors are normally distributed.
- The assumption M.L.R - I through M.L.R - VI are known as **Classic Linear Model(CLM) Assumptions**
- Under these assumptions, OLS is the minimum variance unbiased estimator, the best among all unbiased estimators and not only among those linear in y.

Normal Distribution of OLS Estimators

- Now, under the Classic Linear Model Assumptions, it can be proved that,

$$\hat{\beta}_j \sim Normal[\beta_j, Var(\hat{\beta}_j)]$$

Normal Distribution of OLS Estimators

- Now, under the Classic Linear Model Assumptions, it can be proved that,

$$\hat{\beta}_j \sim Normal[\beta_j, Var(\hat{\beta}_j)]$$

- We can also standardize our beta to make them comparable;

$$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim Normal(0, 1)$$

Normal Distribution of OLS Estimators

- Now, under the Classic Linear Model Assumptions, it can be proved that,

$$\hat{\beta}_j \sim Normal[\beta_j, Var(\hat{\beta}_j)]$$

- We can also standardize our beta to make them comparable;

$$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim Normal(0, 1)$$

- Now, we have all the essential components to do the inference:

Normal Distribution of OLS Estimators

- Now, under the Classic Linear Model Assumptions, it can be proved that,

$$\hat{\beta}_j \sim Normal[\beta_j, Var(\hat{\beta}_j)]$$

- We can also standardize our beta to make them comparable;

$$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim Normal(0, 1)$$

- Now, we have all the essential components to do the inference:
  1. The Mean/Expected Value

Normal Distribution of OLS Estimators

- Now, under the Classic Linear Model Assumptions, it can be proved that,

$$\hat{\beta}_j \sim Normal[\beta_j, Var(\hat{\beta}_j)]$$

- We can also standardize our beta to make them comparable;

$$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim Normal(0, 1)$$

- Now, we have all the essential components to do the inference:
  1. The Mean/Expected Value
  2. The Standard Deviation/Standard Errors

Normal Distribution of OLS Estimators

- Now, under the Classic Linear Model Assumptions, it can be proved that,

$$\hat{\beta}_j \sim Normal[\beta_j, Var(\hat{\beta}_j)]$$

- We can also standardize our beta to make them comparable;

$$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim Normal(0, 1)$$

- Now, we have all the essential components to do the inference:
  1. The Mean/Expected Value
  2. The Standard Deviation/Standard Errors
  3. The Probability Distribution

Challenges to Normality

- We know that a linear combination of independent normal random variables is normal.

Challenges to Normality

- We know that a linear combination of independent normal random variables is normal.
- The usual justification is the central limit theorem. The error is the sum of many unobserved factors, and the sum of many independent factors is approximately normal even if each factor is not. However,

Challenges to Normality

- We know that a linear combination of independent normal random variables is normal.
- The usual justification is the central limit theorem. The error is the sum of many unobserved factors, and the sum of many independent factors is approximately normal even if each factor is not. However,
  1. The factors in the error may have very different distributions, a few of them may dominate, or they may not be independent of each other. The normal approximation can then be poor.

Challenges to Normality

- We know that a linear combination of independent normal random variables is normal.
- The usual justification is the central limit theorem. The error is the sum of many unobserved factors, and the sum of many independent factors is approximately normal even if each factor is not. However,
  1. The factors in the error may have very different distributions, a few of them may dominate, or they may not be independent of each other. The normal approximation can then be poor.
  2. Also, if the factors affect y in a non-additive way, for example multiplicatively, the central limit theorem argument no longer applies.

Challenges to Normality

- We know that a linear combination of independent normal random variables is normal.
- The usual justification is the central limit theorem. The error is the sum of many unobserved factors, and the sum of many independent factors is approximately normal even if each factor is not. However,
  1. The factors in the error may have very different distributions, a few of them may dominate, or they may not be independent of each other. The normal approximation can then be poor.
  2. Also, if the factors affect y in a non-additive way, for example multiplicatively, the central limit theorem argument no longer applies.
- we will try to address these concerns in chapter 5.

Testing Hypotheses About a Single Parameter

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.
- In Econometrics; an important hypothesis is to test if x has an effect on y or not? How can we test this?

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.
- In Econometrics; an important hypothesis is to test if x has an effect on y or not? How can we test this?
-

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0$$

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.
- In Econometrics; an important hypothesis is to test if x has an effect on y or not? How can we test this?
-

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0$$

- The null hypothesis is that x_j has no effect on y whereas the alternate hypothesis is that x_j has some effect on y.

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.
- In Econometrics; an important hypothesis is to test if x has an effect on y or not? How can we test this?
-

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0$$

- The null hypothesis is that x_j has no effect on y whereas the alternate hypothesis is that x_j has some effect on y.
- We either reject the null hypothesis in favor of the alternative or fail to reject it. We cannot;

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.
- In Econometrics; an important hypothesis is to test if x has an effect on y or not? How can we test this?
-

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0$$

- The null hypothesis is that x_j has no effect on y whereas the alternate hypothesis is that x_j has some effect on y.
- We either reject the null hypothesis in favor of the alternative or fail to reject it. We cannot;
  - Accept the Null hypothesis

Hypothesis Testing

- The first inference we can make is with the value of β_j. We can test whether the data are consistent with β_j being equal to a certain value.
- It is important to note here that we are making inference about the population parameter now.
- In Econometrics; an important hypothesis is to test if x has an effect on y or not? How can we test this?
-

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0$$

- The null hypothesis is that x_j has no effect on y whereas the alternate hypothesis is that x_j has some effect on y.
- We either reject the null hypothesis in favor of the alternative or fail to reject it. We cannot;
  - Accept the Null hypothesis
  - Reject the Alternate Hypothesis

t-Test & t-Statistics

Under the CLM assumptions:

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

This is known as t-test statistics/t-ratio/t-calculated. It measures how many standard errors $\hat{\beta}_j$ lies away from your hypothesized value.

- **Rejection Rule:**

$$|t_{n-k-1}| > c$$

Whereas c is known as the critical value.

t-Test & t-Statistics

Under the CLM assumptions:

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

This is known as t-test statistics/t-ratio/t-calculated. It measures how many standard errors $\hat{\beta}_j$ lies away from your hypothesized value.

- **Rejection Rule:**

$$|t_{n-k-1}| > c$$

  Whereas c is known as the critical value.

- The critical value is the number of standard deviation that your hypothesized value can be away given any probability or significance level. It will be different given any significance level.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.
  - We can control for the type I error through significance level. Significance Level of 0.05 means we are allowing the Type-1 error to be 5 percent.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.
  - We can control for the type I error through significance level. Significance Level of 0.05 means we are allowing the Type-1 error to be 5 percent.
  - The critical value depends on our choice of significance level. The higher the significance level, it will be easier to reject null hypothesis.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.
  - We can control for the type I error through significance level. Significance Level of 0.05 means we are allowing the Type-1 error to be 5 percent.
  - The critical value depends on our choice of significance level. The higher the significance level, it will be easier to reject null hypothesis.
  - Therefore, the critical value would be number of standard deviation allowed for our hypothesized value to be away given the probability of Type I error.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.
  - We can control for the type I error through significance level. Significance Level of 0.05 means we are allowing the Type-1 error to be 5 percent.
  - The critical value depends on our choice of significance level. The higher the significance level, it will be easier to reject null hypothesis.
  - Therefore, the critical value would be number of standard deviation allowed for our hypothesized value to be away given the probability of Type I error.
- **Type II Error** - Unable to reject the Null hypothesis, When it is actually false.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.
  - We can control for the type I error through significance level. Significance Level of 0.05 means we are allowing the Type-1 error to be 5 percent.
  - The critical value depends on our choice of significance level. The higher the significance level, it will be easier to reject null hypothesis.
  - Therefore, the critical value would be number of standard deviation allowed for our hypothesized value to be away given the probability of Type I error.
- **Type II Error** - Unable to reject the Null hypothesis, When it is actually false.
  - We can control for the type II error through number of observation and the variance of OLS estimators.

Error in Hypothesis Testing

- **Type I Error** - Rejecting the Null Hypothesis, when it is actually true.
  - We can control for the type I error through significance level. Significance Level of 0.05 means we are allowing the Type-1 error to be 5 percent.
  - The critical value depends on our choice of significance level. The higher the significance level, it will be easier to reject null hypothesis.
  - Therefore, the critical value would be number of standard deviation allowed for our hypothesized value to be away given the probability of Type I error.
- **Type II Error** - Unable to reject the Null hypothesis, When it is actually false.
  - We can control for the type II error through number of observation and the variance of OLS estimators.
  - The higher the number of observations or the lower variance of OLS estimators, the lower the probability of Type II - Error.

Two-Tailed Test of Significance

- $$H_0 : \beta_j = 0$$
$$H_A : \beta_j \neq 0$$

Two-Tailed Test of Significance

- $$H_0 : \beta_j = 0$$
$$H_A : \beta_j \neq 0$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

Two-Tailed Test of Significance

- $$H_0 : \beta_j = 0$$
$$H_A : \beta_j \neq 0$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

- The critical value of t would come from the table at given significance level.

Two-Tailed Test of Significance

- $$H_0 : \beta_j = 0$$
$$H_A : \beta_j \neq 0$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

- The critical value of t would come from the table at given significance level.
- we will reject the null hypothesis at

$$|t_{n-k-1}| > c$$

Two-Tailed Test of Significance

- $$H_0 : \beta_j = 0$$
  $$H_A : \beta_j \neq 0$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

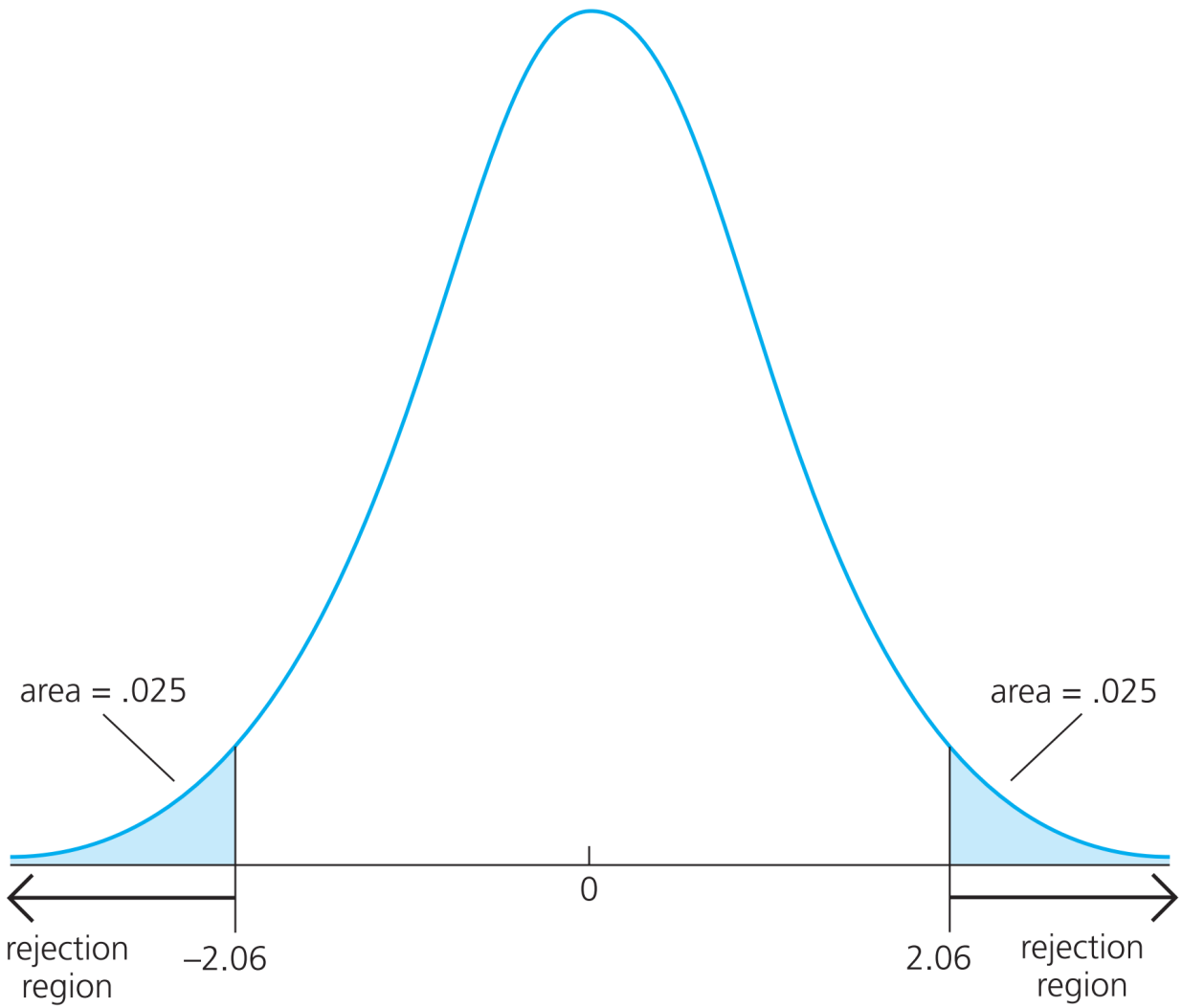
- The critical value of t would come from the table at given significance level.
- we will reject the null hypothesis at

$$|t_{n-k-1}| > c$$

- If we reject the null hypothesis, then we would say that we reject the null hypothesis of insignificance and conclude that x_j is statistically significant or x_j has a statistically significant effect on y.

Two-Tailed Test of Significance

Figure 4.4 5% rejection rule for the alternative H_1: $\beta_j \neq 0$ with 25 *df*.



Two-Tailed Test of Significance

Example:

EXAMPLE 4.3 **Determinants of College GPA**

We use the data in GPA1 to estimate a model explaining college GPA (*colGPA*), with the average number of lectures missed per week (*skipped*) as an additional explanatory variable. The estimated model is

$$\widehat{colGPA} = 1.39 + .412\, hsGPA + .015\, ACT - .083\, skipped$$
$$\qquad\qquad (.33)\quad (.094)\qquad\quad (.011)\qquad (.026)$$
$$n = 141,\ R^2 = .234.$$

- $$t_{141-3-1=137} = \frac{0.412}{0.094} = 4.38$$

Two-Tailed Test of Significance

Example:

EXAMPLE 4.3 **Determinants of College GPA**

We use the data in GPA1 to estimate a model explaining college GPA (*colGPA*), with the average number of lectures missed per week (*skipped*) as an additional explanatory variable. The estimated model is

$$\widehat{colGPA} = 1.39 + .412\, hsGPA + .015\, ACT - .083\, skipped$$
$$(.33) \quad (.094) \qquad (.011) \qquad (.026)$$
$$n = 141,\ R^2 = .234.$$

- $$t_{141-3-1=137} = \frac{0.412}{0.094} = 4.38$$
- The critical value at df $= 137$ and $\alpha = 0.05$ & 0.01 is approximately 1.96 & 2.58 respectively, using the large sample values (exact values 1.98 & 2.61). (we will see the tabulated value at 0.025 and 0.005)

Two-Tailed Test of Significance

Example:

EXAMPLE 4.3 **Determinants of College GPA**

We use the data in GPA1 to estimate a model explaining college GPA (*colGPA*), with the average number of lectures missed per week (*skipped*) as an additional explanatory variable. The estimated model is

$$\widehat{colGPA} = 1.39 + .412\, hsGPA + .015\, ACT - .083\, skipped$$
$$(.33) \quad (.094) \qquad (.011) \qquad (.026)$$
$$n = 141,\ R^2 = .234.$$

- $$t_{141-3-1=137} = \frac{0.412}{0.094} = 4.38$$
- The critical value at df $= 137$ and $\alpha = 0.05$ & 0.01 is approximately 1.96 & 2.58 respectively, using the large sample values (exact values 1.98 & 2.61). (we will see the tabulated value at 0.025 and 0.005)
- As $t_{137} = 4.38 > 1.96\&2.58$, we reject the null hypothesis of insignificance and conclude that High School GPA has a statistically significant effect on the college GPA.

Two-Tailed Test of a Certain Value

- $$H_0 : \beta_j = a$$
$$H_A : \beta_j \neq a$$

Two-Tailed Test of a Certain Value

- $$H_0 : \beta_j = a$$
$$H_A : \beta_j \neq a$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

Two-Tailed Test of a Certain Value

- $$H_0 : \beta_j = a$$
$$H_A : \beta_j \neq a$$
- The t-test statistics would be;
$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$
- The critical value of t would come from the table at given significance level. (The tabulated value at $\frac{\alpha}{2}$)

Two-Tailed Test of a Certain Value

- $$H_0 : \beta_j = a$$
$$H_A : \beta_j \neq a$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

- The critical value of t would come from the table at given significance level. (The tabulated value at $\frac{\alpha}{2}$)
- we will reject the null hypothesis at

$$|t_{n-k-1}| > c$$

Two-Tailed Test of a Certain Value

- $$H_0 : \beta_j = a$$
$$H_A : \beta_j \neq a$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

- The critical value of t would come from the table at given significance level. (The tabulated value at $\frac{\alpha}{2}$)
- we will reject the null hypothesis at

$$|t_{n-k-1}| > c$$

- If we reject the null hypothesis, we conclude that β_j is statistically different from a at the chosen significance level.

Two-Tailed Test of a Certain Value

Example:

EXAMPLE 4.3 **Determinants of College GPA**

We use the data in GPA1 to estimate a model explaining college GPA (*colGPA*), with the average number of lectures missed per week (*skipped*) as an additional explanatory variable. The estimated model is

$$\widehat{colGPA} = 1.39 + .412\, hsGPA + .015\, ACT - .083\, skipped$$
$$(.33) \quad (.094) \qquad (.011) \qquad (.026)$$
$$n = 141,\ R^2 = .234.$$

Test if $\beta_j = 0.2$

- $$t_{141-3-1=137} = \frac{0.412 - 0.2}{0.094} = 2.26$$

Two-Tailed Test of a Certain Value

Example:

EXAMPLE 4.3 **Determinants of College GPA**

We use the data in GPA1 to estimate a model explaining college GPA (*colGPA*), with the average number of lectures missed per week (*skipped*) as an additional explanatory variable. The estimated model is

$$\widehat{colGPA} = 1.39 + .412\, hsGPA + .015\, ACT - .083\, skipped$$
$$(.33) \quad (.094) \qquad (.011) \qquad (.026)$$
$$n = 141,\ R^2 = .234.$$

Test if $\beta_j = 0.2$

- $$t_{141-3-1=137} = \frac{0.412 - 0.2}{0.094} = 2.26$$
- The critical value at df $= 137$ and $\alpha = 0.05$ & 0.01 is approximately 1.96 & 2.58 respectively, using the large sample values (exact values 1.98 & 2.61). (we will see the tabulated value at 0.025 and 0.005)

Two-Tailed Test of a Certain Value

Example:

EXAMPLE 4.3 **Determinants of College GPA**

We use the data in GPA1 to estimate a model explaining college GPA (*colGPA*), with the average number of lectures missed per week (*skipped*) as an additional explanatory variable. The estimated model is

$$\widehat{colGPA} = 1.39 + .412\, hsGPA + .015\, ACT - .083\, skipped$$
$$(.33) \quad (.094) \qquad (.011) \qquad (.026)$$
$$n = 141,\ R^2 = .234.$$

Test if $\beta_j = 0.2$

- $$t_{141-3-1=137} = \frac{0.412 - 0.2}{0.094} = 2.26$$
- The critical value at df $= 137$ and $\alpha = 0.05$ & 0.01 is approximately 1.96 & 2.58 respectively, using the large sample values (exact values 1.98 & 2.61). (we will see the tabulated value at 0.025 and 0.005)
- As $t_{137} = 2.26 > 1.96$, we reject $H_0 : \beta_j = 0.2$ at the 5% level and conclude that the effect of High School GPA differs from 0.2. However, as $2.26 < 2.58$, we fail to reject it at the 1% level.

One-Tailed Test

- $$H_0 : \beta_j \leq a$$
$$H_A : \beta_j > a$$

One-Tailed Test

- $$H_0 : \beta_j \leq a$$
$$H_A : \beta_j > a$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

One-Tailed Test

- $$H_0 : \beta_j \leq a$$
$$H_A : \beta_j > a$$
- The t-test statistics would be;

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

- The critical value of t would come from the table at given significance level. (the tabulated value is taken at α)

One-Tailed Test

- $$H_0 : \beta_j \leq a$$
$$H_A : \beta_j > a$$
- The t-test statistics would be;

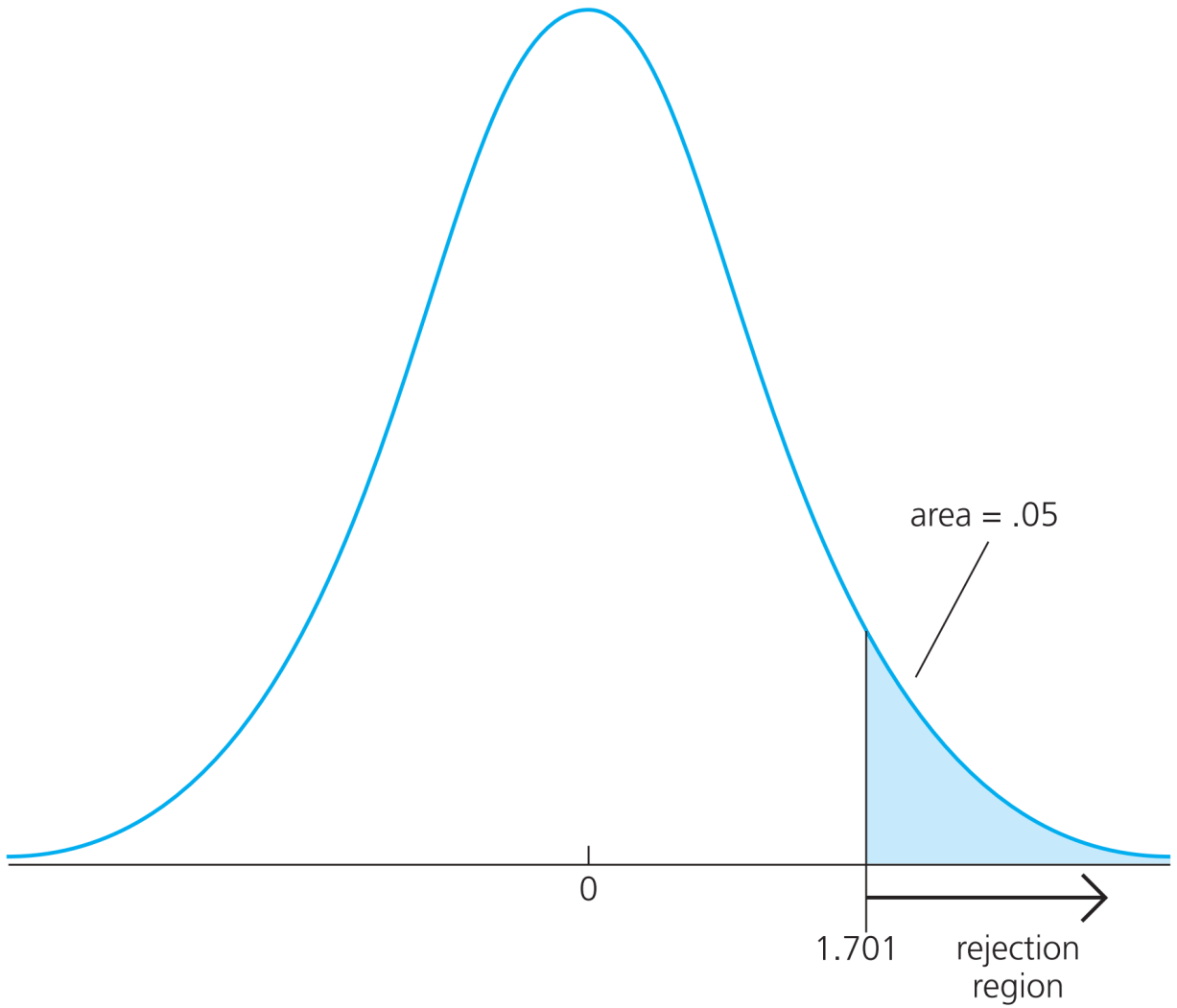
$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$$

- The critical value of t would come from the table at given significance level. (the tabulated value is taken at α)
- we will reject the null hypothesis at

$$t_{n-k-1} > c$$

One Tailed Test

Figure 4.2 5% rejection rule for the alternative H_1: $\beta_j > 0$ with 28 *df*.



Computing P-Values for t-Test

- As we have noticed that the Estimators can be statistically significant at one significance level but insignificant at a different significance level.

Computing P-Values for t-Test

- As we have noticed that the Estimators can be statistically significant at one significance level but insignificant at a different significance level.
- So, instead of testing our hypothesis at different significance levels, we can compute the p-value which is the smallest significance level at which the null hypothesis can be rejected.

$$P(|T| > |t|)$$

Computing P-Values for t-Test

- As we have noticed that the Estimators can be statistically significant at one significance level but insignificant at a different significance level.
- So, instead of testing our hypothesis at different significance levels, we can compute the p-value which is the smallest significance level at which the null hypothesis can be rejected.

$$P(|T| > |t|)$$

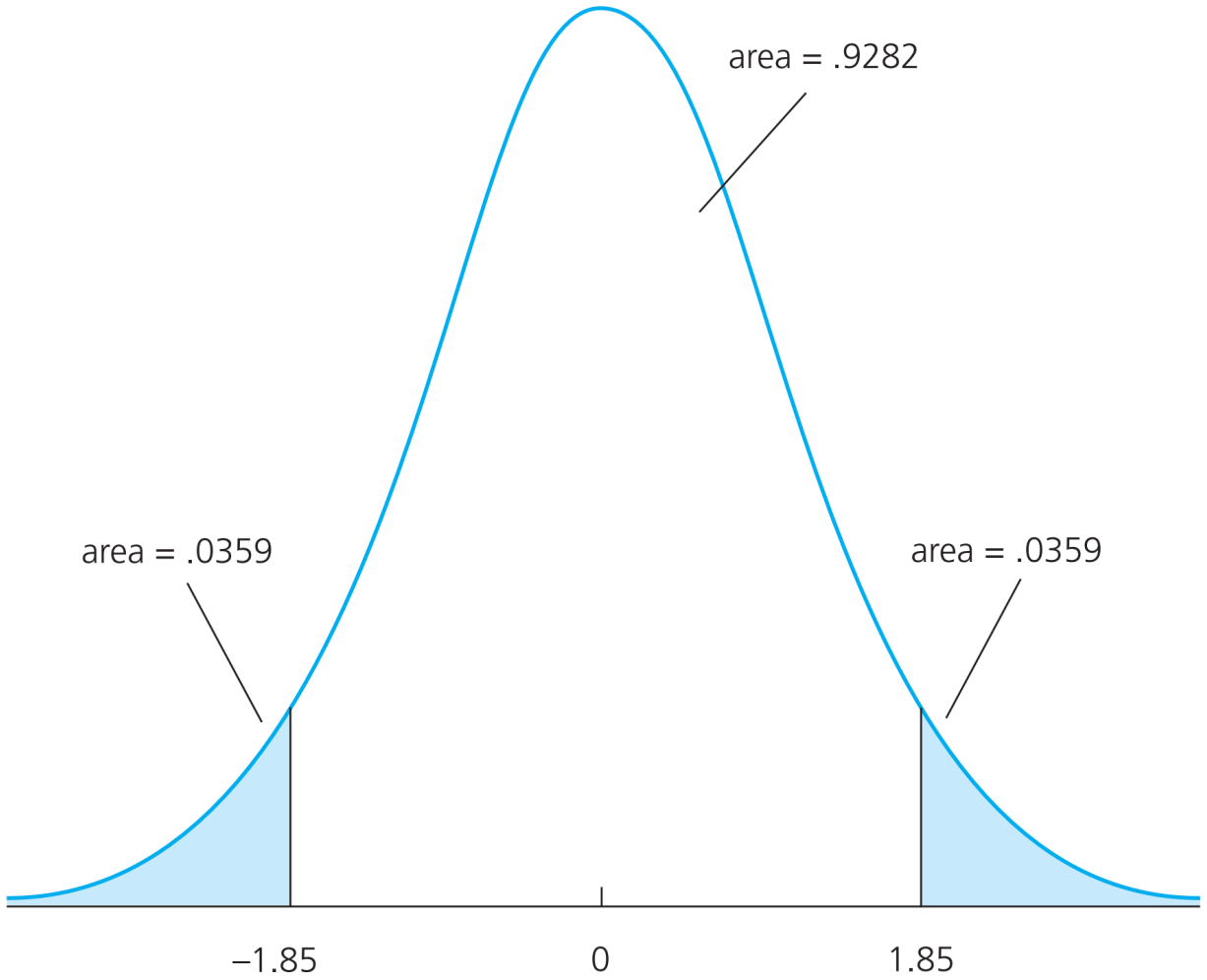
- Formally, the p-value is the probability that a t distributed random variable, under H_0, is larger in absolute value than the observed test statistic.

$$P(|T| > |t|)$$

Where T is a random variable following the t_{n-k-1} distribution under H_0 & t is the observed test statistic.

Computing P-Values for t-Test

Figure 4.6 Obtaining the *p*-value against a two-sided alternative, when $t = 1.85$ and $df = 40$.



Example

```
. regress price mpg rep78 headroom
```

Source	SS	df	MS		
				Number of obs	= 69
				F(3, 65)	= 7.51
Model	148497605	3	49499201.8	Prob > F	= 0.0002
Residual	428299354	65	6589220.82	R-squared	= 0.2575
				Adj R-squared	= 0.2232
Total	576796959	68	8482308.22	Root MSE	= 2566.9

price	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
mpg	-289.3462	62.53921	-4.63	0.000	-414.2456	-164.4467
rep78	670.8971	343.5213	1.95	0.055	-15.16242	1356.957
headroom	-300.0293	398.0516	-0.75	0.454	-1094.993	494.9346
_cons	10921.33	2153.003	5.07	0.000	6621.487	15221.17

Different Kinds of Significances

1. So far, we have covered the statistical significance of estimators. Sometimes, you can get fixated on the statistical significance and ignore the size and direction of estimator.

Different Kinds of Significances

1. So far, we have covered the statistical significance of estimators. Sometimes, you can get fixated on the statistical significance and ignore the size and direction of estimator.
2. **Economic Significance** - The economic significance is about the size and sign of the estimate. Is the effect large enough to matter, and is its direction consistent with theory?

Different Kinds of Significances

1. So far, we have covered the statistical significance of estimators. Sometimes, you can get fixated on the statistical significance and ignore the size and direction of estimator.
2. **Economic Significance** - The economic significance is about the size and sign of the estimate. Is the effect large enough to matter, and is its direction consistent with theory?
3. **Practical Significance** - Wooldridge uses this term interchangeably with economic significance. Is a statistically significant but tiny effect more meaningful than a large but imprecisely estimated one?

Different Kinds of Significances

1. So far, we have covered the statistical significance of estimators. Sometimes, you can get fixated on the statistical significance and ignore the size and direction of estimator.
2. **Economic Significance** - The economic significance is about the size and sign of the estimate. Is the effect large enough to matter, and is its direction consistent with theory?
3. **Practical Significance** - Wooldridge uses this term interchangeably with economic significance. Is a statistically significant but tiny effect more meaningful than a large but imprecisely estimated one?
4. DO NOT forget that you are economist and not statistician. Always look beyond just technical stuff.

Confidence Intervals

Confidence Intervals

- Under the CLM assumptions, we can easily construct confidence interval for OLS parameters.

Confidence Intervals

- Under the CLM assumptions, we can easily construct confidence interval for OLS parameters.
- At a given significance level α, confidence interval can be defined as:

$$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$

where c is the upper $\frac{\alpha}{2}$ critical value, the $100(1-\frac{\alpha}{2})$th percentile of the t_{n-k-1} distribution.

Confidence Intervals

- Under the CLM assumptions, we can easily construct confidence interval for OLS parameters.
- At a given significance level α, confidence interval can be defined as:

$$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$

where c is the upper $\frac{\alpha}{2}$ critical value, the $100(1-\frac{\alpha}{2})$th percentile of the t_{n-k-1} distribution.

- $$\hat{\beta}_j - c.se(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + c.se(\hat{\beta}_j)$$

Confidence Intervals

- Under the CLM assumptions, we can easily construct confidence interval for OLS parameters.
- At a given significance level α, confidence interval can be defined as:

$$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$

where c is the upper $\frac{\alpha}{2}$ critical value, the $100(1-\frac{\alpha}{2})$th percentile of the t_{n-k-1} distribution.

- $$\hat{\beta}_j - c.se(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + c.se(\hat{\beta}_j)$$
- If a random sample is drawn 100 times and a confidence interval is built each time, about $(1-\alpha) \times 100$ of these intervals would contain the population parameter.

Confidence Intervals

- Under the CLM assumptions, we can easily construct confidence interval for OLS parameters.
- At a given significance level α, confidence interval can be defined as:

$$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$

where c is the upper $\frac{\alpha}{2}$ critical value, the $100(1-\frac{\alpha}{2})$th percentile of the t_{n-k-1} distribution.

- $$\hat{\beta}_j - c.se(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + c.se(\hat{\beta}_j)$$
- If a random sample is drawn 100 times and a confidence interval is built each time, about $(1-\alpha) \times 100$ of these intervals would contain the population parameter.
- It is important to note that confidence intervals are as good as the validity of the CLM assumptions.

Example:

$$\widehat{\log(rd)} = -4.38 + 1.084\log(sales) + .0217\,profmarg$$
$$(.47) \quad (.060) \qquad\qquad (.0128)$$
$$n = 32, R^2 = .918.$$

- Find the 95 percent Confidence Interval of the slope estimator of sales. given c = 1.96

Example:

$$\widehat{\log(rd)} = -4.38 + 1.084\log(sales) + .0217\,profmarg$$
$$(.47) \quad (.060) \qquad\qquad (.0128)$$
$$n = 32, R^2 = .918.$$

- Find the 95 percent Confidence Interval of the slope estimator of sales. given c = 1.96
-

$$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$

Example:

$$\widehat{\log(rd)} = -4.38 + 1.084\log(sales) + .0217\,profmarg$$
$$(.47) \quad (.060) \qquad\qquad (.0128)$$
$$n = 32, R^2 = .918.$$

- Find the 95 percent Confidence Interval of the slope estimator of sales. given c = 1.96
- $$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$
- $$1.084 \pm 1.96 \times 0.06$$

Example:

$$\widehat{\log(rd)} = -4.38 + 1.084\log(sales) + .0217\,profmarg$$
$$(.47) \quad (.060) \qquad\qquad (.0128)$$
$$n = 32, R^2 = .918.$$

- Find the 95 percent Confidence Interval of the slope estimator of sales. given c = 1.96
- $$\hat{\beta}_j \pm c.se(\hat{\beta}_j)$$
- $$1.084 \pm 1.96 \times 0.06$$
- $$0.9664 \leq \beta_{sales} \leq 1.2016$$

Testing Hypotheses about a Single Linear Combination of Parameters

Testing Hypotheses about a Single Linear Combination of Parameters

- Consider the following model:

$$log(wage) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 exper + \mu$$

where jc is number of years attending a two-year college.

univ is number of years at a four-year college.

exper is months in the workforce.

Testing Hypotheses about a Single Linear Combination of Parameters

- Consider the following model:

$$log(wage) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 exper + \mu$$

where jc is number of years attending a two-year college.

univ is number of years at a four-year college.

exper is months in the workforce.

- $$H_0 : \beta_1 = \beta_2$$
$$H_A : \beta_1 \neq \beta_2$$

Testing Hypotheses about a Single Linear Combination of Parameters

- Consider the following model:

$$log(wage) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 exper + \mu$$

where jc is number of years attending a two-year college.

univ is number of years at a four-year college.

exper is months in the workforce.

- $$H_0 : \beta_1 = \beta_2$$
$$H_A : \beta_1 \neq \beta_2$$

- It can also be written as:

$$H_0 : \beta_1 - \beta_2 = 0$$
$$H_A : \beta_1 - \beta_2 \neq 0$$

Testing Hypotheses about a Single Linear Combination of Parameters

- $$t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{se(\hat{\beta}_1 - \hat{\beta}_2)} \sim t_{n-k-1}$$

Testing Hypotheses about a Single Linear Combination of Parameters

- $$t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{se(\hat{\beta}_1 - \hat{\beta}_2)} \sim t_{n-k-1}$$
- where;

$$Var(\hat{\beta}_1 - \hat{\beta}_2) = Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)$$

Testing Hypotheses about a Single Linear Combination of Parameters

- $$t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{se(\hat{\beta}_1 - \hat{\beta}_2)} \sim t_{n-k-1}$$
- where;

$$Var(\hat{\beta}_1 - \hat{\beta}_2) = Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)$$

- So,

$$se(\hat{\beta}_1 - \hat{\beta}_2) = \sqrt{Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)}$$

Testing Hypotheses about a Single Linear Combination of Parameters

- $$t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{se(\hat{\beta}_1 - \hat{\beta}_2)} \sim t_{n-k-1}$$
- where;

$$Var(\hat{\beta}_1 - \hat{\beta}_2) = Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)$$

- So,

$$se(\hat{\beta}_1 - \hat{\beta}_2) = \sqrt{Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)}$$

- Think about any one-tail test hypothesis involving single linear combination of parameter.

Testing Hypotheses about a Multiple Linear Combination of Parameters

Testing Exclusion Restrictions: The F-Test

- Now, we want to test more than one hypothesis together. For Example:

$$\beta_3 = 0, \beta_4 = 0, \beta_5 = 0$$

Testing Exclusion Restrictions: The F-Test

- Now, we want to test more than one hypothesis together. For Example:

$$\beta_3 = 0, \beta_4 = 0, \beta_5 = 0$$

- We can test them separately using t-test. However, the joint testing is different from individual testing.

Testing Exclusion Restrictions: The F-Test

- Now, we want to test more than one hypothesis together. For Example:

$$\beta_3 = 0, \beta_4 = 0, \beta_5 = 0$$

- We can test them separately using t-test. However, the joint testing is different from individual testing.
- Let's look at the following model: $log(salary) = \beta_0 + \beta_1 years + \beta_2 gamesyr + \beta_3 bavg + \beta_4 hrunsyr + \beta_5 rbisyr + \mu$ where bavg, hrunsyr, rbisyr are their performance indicators in the game.

Testing Exclusion Restrictions: The F-Test

- Now, we want to test more than one hypothesis together. For Example:

$$\beta_3 = 0, \beta_4 = 0, \beta_5 = 0$$

- We can test them separately using t-test. However, the joint testing is different from individual testing.
- Let's look at the following model: $log(salary) = \beta_0 + \beta_1 years + \beta_2 gamesyr + \beta_3 bavg + \beta_4 hrunsyr + \beta_5 rbisyr + \mu$ where bavg, hrunsyr, rbisyr are their performance indicators in the game.
- When we want to test the effect of one performance indicator, we will do the t-testing. However, if we want to test that does performance affect the salary?, we have to do the joint testing.

Testing Exclusion Restrictions: The F-Test

- In joint testing, we call the no effect of multiple variables as **Exclusion Restrictions**. It means that multiple variables can be excluded from the model.

Testing Exclusion Restrictions: The F-Test

- In joint testing, we call the no effect of multiple variables as **Exclusion Restrictions**. It means that multiple variables can be excluded from the model.
- $$H_0 : \beta_3 = \beta_4 = \beta_5 = 0$$

  $$H_A : Not(H_0)$$

  The alternative hypothesis is true if at least one of these parameters is not zero.

Testing Exclusion Restrictions: The F-Test

- In joint testing, we call the no effect of multiple variables as **Exclusion Restrictions**. It means that multiple variables can be excluded from the model.
- $$H_0 : \beta_3 = \beta_4 = \beta_5 = 0$$

  $$H_A : Not(H_0)$$

  The alternative hypothesis is true if at least one of these parameters is not zero.
- To do this test, we will estimate two models.

Testing Exclusion Restrictions: The F-Test

- In joint testing, we call the no effect of multiple variables as **Exclusion Restrictions**. It means that multiple variables can be excluded from the model.
- $$H_0 : \beta_3 = \beta_4 = \beta_5 = 0$$
  $$H_A : Not(H_0)$$
  The alternative hypothesis is true if at least one of these parameters is not zero.
- To do this test, we will estimate two models.
  1. UnRestricted Model - Same as the last slide (Original Model)

Testing Exclusion Restrictions: The F-Test

- In joint testing, we call the no effect of multiple variables as **Exclusion Restrictions**. It means that multiple variables can be excluded from the model.
- $$H_0 : \beta_3 = \beta_4 = \beta_5 = 0$$
  $$H_A : Not(H_0)$$
  The alternative hypothesis is true if at least one of these parameters is not zero.
- To do this test, we will estimate two models.
  1. UnRestricted Model - Same as the last slide (Original Model)
  2. Restricted Model - Where Null Hypothesis is True
     $log(salary) = \beta_0 + \beta_1 years + \beta_2 gamesyr + \mu$

Testing Exclusion Restrictions: The F-Test

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim \mathbf{F}_{q,n-k-1}$$

  where F is the test statistics (F-Statistics)

  SSR is Residual Sum of Squares,

  q is the number of restrictions (Numerator Degree of Freedom $df_r - df_{ur}$)

Testing Exclusion Restrictions: The F-Test

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim \mathbf{F}_{q,n-k-1}$$

  where F is the test statistics (F-Statistics)

  SSR is Residual Sum of Squares,

  q is the number of restrictions (Numerator Degree of Freedom $df_r - df_{ur}$)
- The F-statistic measures the relative increase in SSR from the unrestricted to the restricted model.

Testing Exclusion Restrictions: The F-Test

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim \mathbf{F}_{q,n-k-1}$$

  where F is the test statistics (F-Statistics)

  SSR is Residual Sum of Squares,

  q is the number of restrictions (Numerator Degree of Freedom $df_r - df_{ur}$)
- The F-statistic measures the relative increase in SSR from the unrestricted to the restricted model.
- We will look at whether the increase in SSR in going from unrestricted model to restricted model is large enough to get the rejection.

Testing Exclusion Restrictions: The F-Test

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim \mathbf{F}_{q,n-k-1}$$

  where F is the test statistics (F-Statistics)

  SSR is Residual Sum of Squares,

  q is the number of restrictions (Numerator Degree of Freedom $df_r - df_{ur}$)
- The F-statistic measures the relative increase in SSR from the unrestricted to the restricted model.
- We will look at whether the increase in SSR in going from unrestricted model to restricted model is large enough to get the rejection.
- we will reject the H_0 if $F > c$. If H_0 is rejected then we say x_3, x_4, x_5 are jointly statistically significant at a certain significance level. If we cannot reject the null hypothesis, the variables are jointly insignificant and in most of the cases we drop them.

Testing Exclusion Restrictions: The F-Test

- As SSR is related to R-Squared. We can change our formulae in terms of R-Squared as well.

Testing Exclusion Restrictions: The F-Test

- As SSR is related to R-Squared. We can change our formulae in terms of R-Squared as well.
- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k - 1)} \sim \mathbf{F}_{q,n-k-1}$$

Testing Exclusion Restrictions: The F-Test

- As SSR is related to R-Squared. We can change our formulae in terms of R-Squared as well.
- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k - 1)} \sim \mathbf{F}_{q,n-k-1}$$
- Note, the signs are reversed now as there is an inverse relationship between SSR and R-Squared.

Testing Exclusion Restrictions: The F-Test

- As SSR is related to R-Squared. We can change our formulae in terms of R-Squared as well.
- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k - 1)} \sim \mathbf{F}_{q,n-k-1}$$
- Note, the signs are reversed now as there is an inverse relationship between SSR and R-Squared.
- Also, Note that the F-Statistics can never be negative as $R^2_{ur} \geq R^2_r$. However, we need to see how much the R-squared go down after the restrictions.

Testing Exclusion Restrictions: The F-Test

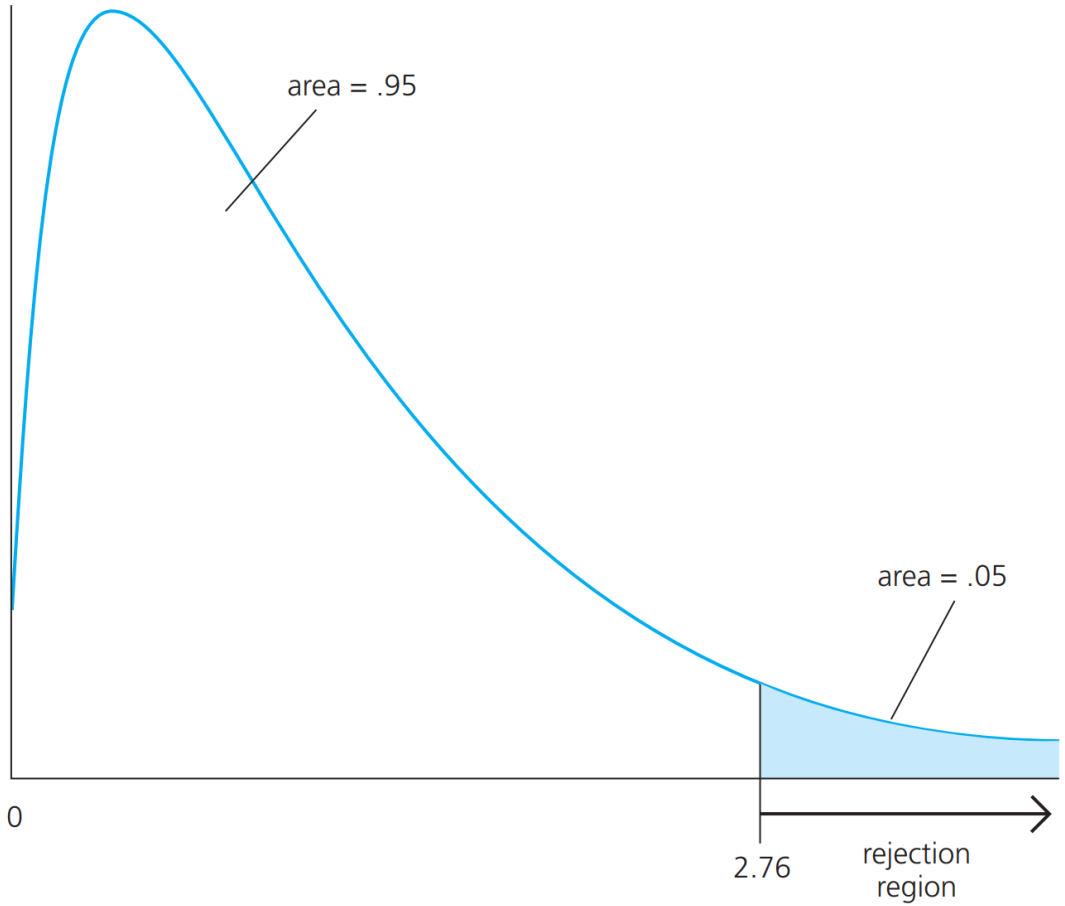
- As SSR is related to R-Squared. We can change our formulae in terms of R-Squared as well.
- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k - 1)} \sim \mathbf{F}_{q,n-k-1}$$
- Note, the signs are reversed now as there is an inverse relationship between SSR and R-Squared.
- Also, Note that the F-Statistics can never be negative as $R^2_{ur} \geq R^2_r$. However, we need to see how much the R-squared go down after the restrictions.
- Under H_0, the numerator and denominator are independent chi-square random variables, each divided by its degrees of freedom, and such a ratio follows an F-distribution.

Testing Exclusion Restrictions: The F-Test

- As SSR is related to R-Squared. We can change our formulae in terms of R-Squared as well.
- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k - 1)} \sim \mathbf{F}_{q,n-k-1}$$
- Note, the signs are reversed now as there is an inverse relationship between SSR and R-Squared.
- Also, Note that the F-Statistics can never be negative as $R^2_{ur} \geq R^2_r$. However, we need to see how much the R-squared go down after the restrictions.
- Under H_0, the numerator and denominator are independent chi-square random variables, each divided by its degrees of freedom, and such a ratio follows an F-distribution.
- Lastly, It can be shown that the F-distribution with one restriction is equivalent to the square of relevant t-distribution $t^2_{n-k-1} = F_{1,n-k-1}$

Shape of F-Distribution

Figure 4.7 The 5% critical value and rejection region in an $F_{3,60}$ distribution.



Example

- Un-Restricted Model

$$\widehat{\log(salary)} = 11.19 + .0689\ years + .0126\ gamesyr$$
$$(0.29)\ \ (.0121)\ \ \ \ \ (.0026)$$
$$+ .00098\ bavg + .0144\ hrunsyr + .0108\ rbisyr \qquad [4.31]$$
$$(.00110)\ \ \ \ \ (.0161)\ \ \ \ \ (.0072)$$
$$n = 353, \text{SSR} = 183.186, R^2 = .6278,$$

Example

- Un-Restricted Model

$$\widehat{\log(salary)} = 11.19 + .0689\ years + .0126\ gamesyr$$
$$(0.29)\ \ (.0121)\ \ \ \ \ (.0026)$$
$$+ .00098\ bavg + .0144\ hrunsyr + .0108\ rbisyr \qquad [4.31]$$
$$(.00110)\ \ \ \ \ (.0161)\ \ \ \ \ (.0072)$$
$$n = 353, \text{SSR} = 183.186, R^2 = .6278,$$

- Restricted Model

$$\widehat{\log(salary)} = 11.22 + .0713\ years + .0202\ gamesyr$$
$$(.11)\ \ (.0125)\ \ \ \ \ (.0013)$$
$$n = 353, \text{SSR} = 198.311, R^2 = .5971. \qquad [4.33]$$

Example

- Un-Restricted Model

$$\widehat{\log(salary)} = 11.19 + .0689\ years + .0126\ gamesyr + .00098\ bavg + .0144\ hrunsyr + .0108\ rbisyr \quad [4.31]$$
$$(0.29)\ \ (.0121)\ \ (.0026)\ \ (.00110)\ \ (.0161)\ \ (.0072)$$
$$n = 353,\ \text{SSR} = 183.186,\ R^2 = .6278,$$

- Restricted Model

$$\widehat{\log(salary)} = 11.22 + .0713\ years + .0202\ gamesyr \quad [4.33]$$
$$(.11)\ \ (.0125)\ \ (.0013)$$
$$n = 353,\ \text{SSR} = 198.311,\ R^2 = .5971.$$

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} = \frac{(198.311 - 183.186)/3}{183.186/347} = 9.55 \sim \mathbf{F}_{q,n-k-1}$$

Example

- Un-Restricted Model

$$\widehat{\log(salary)} = 11.19 + .0689\ years + .0126\ gamesyr + .00098\ bavg + .0144\ hrunsyr + .0108\ rbisyr \quad [4.31]$$
$$(0.29)\ \ (.0121)\ \ (.0026)\ \ (.00110)\ \ (.0161)\ \ (.0072)$$
$$n = 353,\ \text{SSR} = 183.186,\ R^2 = .6278,$$

- Restricted Model

$$\widehat{\log(salary)} = 11.22 + .0713\ years + .0202\ gamesyr \quad [4.33]$$
$$(.11)\ \ (.0125)\ \ (.0013)$$
$$n = 353,\ \text{SSR} = 198.311,\ R^2 = .5971.$$

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} = \frac{(198.311 - 183.186)/3}{183.186/347} = 9.55 \sim \mathbf{F}_{q,n-k-1}$$

- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1-R^2_{ur})/(n-k-1)} = \frac{(0.6278 - 0.5971)/3}{(1-0.6278)/347} = 9.55 \sim \mathbf{F}_{q,n-k-1}$$

Example

- Un-Restricted Model

$$\widehat{\log(salary)} = 11.19 + .0689\ years + .0126\ gamesyr + .00098\ bavg + .0144\ hrunsyr + .0108\ rbisyr \quad [4.31]$$
$$(0.29)\ \ (.0121)\ \ (.0026)\ \ (.00110)\ \ (.0161)\ \ (.0072)$$
$$n = 353,\ \text{SSR} = 183.186,\ R^2 = .6278,$$

- Restricted Model

$$\widehat{\log(salary)} = 11.22 + .0713\ years + .0202\ gamesyr \quad [4.33]$$
$$(.11)\ \ (.0125)\ \ (.0013)$$
$$n = 353,\ \text{SSR} = 198.311,\ R^2 = .5971.$$

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} = \frac{(198.311 - 183.186)/3}{183.186/347} = 9.55 \sim \mathbf{F}_{q,n-k-1}$$
- $$F = \frac{(R^2_{ur} - R^2_r)/q}{(1-R^2_{ur})/(n-k-1)} = \frac{(0.6278 - 0.5971)/3}{(1-0.6278)/347} = 9.55 \sim \mathbf{F}_{q,n-k-1}$$
- The 5% critical value of $F_{3,347}$ is about 2.63 (3.84 at 1%). As $F > c$, we reject the null hypothesis and conclude that the performance indicators are jointly significant.

The F-Statistics of Overall Significance

- Using the same idea, we can test if our model helps to explain y. In other words, does average of y is a better estimator than the model.

The F-Statistics of Overall Significance

- Using the same idea, we can test if our model helps to explain y. In other words, does average of y is a better estimator than the model.
- $H_0 : x_1, x_2, x_3, ..., x_k$ do not help to explain y $\beta_1 = \beta_2 = ... = \beta_k = 0$

The F-Statistics of Overall Significance

- Using the same idea, we can test if our model helps to explain y. In other words, does average of y is a better estimator than the model.
- $H_0 : x_1, x_2, x_3, ..., x_k$ do not help to explain y $\beta_1 = \beta_2 = ... = \beta_k = 0$
- Un-Restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \mu$$

The F-Statistics of Overall Significance

- Using the same idea, we can test if our model helps to explain y. In other words, does average of y is a better estimator than the model.
- $H_0 : x_1, x_2, x_3, ..., x_k$ do not help to explain y $\beta_1 = \beta_2 = ... = \beta_k = 0$
- Un-Restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \mu$$

- Restricted Model:

$$y = \beta_0 \qquad R^2 = 0, k \text{ restrictions}$$

The F-Statistics of Overall Significance

- Using the same idea, we can test if our model helps to explain y. In other words, does average of y is a better estimator than the model.
- $H_0 : x_1, x_2, x_3, ..., x_k$ do not help to explain y $\beta_1 = \beta_2 = ... = \beta_k = 0$
- Un-Restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \mu$$

- Restricted Model:

$$y = \beta_0 \qquad R^2 = 0, k \textit{ restrictions}$$

-

$$F = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k - 1)} = \frac{(R^2)/k}{(1 - R^2)/(n - k - 1)} \sim \mathbf{F}_{q,n-k-1}$$

The F-Statistics of Overall Significance

- Now, if $F > c$, we will reject the null hypothesis and say that the model is jointly statistically significant.

The F-Statistics of Overall Significance

- Now, if $F > c$, we will reject the null hypothesis and say that the model is jointly statistically significant.
- If we are unable to reject the null hypothesis, we will say that the model is jointly statistically insignificant. In this case the average of y is a better estimator.

The F-Statistics of Overall Significance

- Now, if $F > c$, we will reject the null hypothesis and say that the model is jointly statistically significant.
- If we are unable to reject the null hypothesis, we will say that the model is jointly statistically insignificant. In this case the average of y is a better estimator.
- Example:

$$\widehat{\log(price)} = \underset{(.570)}{.264} + \underset{(.151)}{1.043}\log(assess) + \underset{(.0386)}{.0074}\log(lotsize) - \underset{(.1384)}{.1032}\log(sqrft) + \underset{(.0221)}{.0338}\,bdrms$$

$$n = 88,\ \text{SSR} = 1.822,\ R^2 = .773.$$

The F-Statistics of Overall Significance

- Now, if $F > c$, we will reject the null hypothesis and say that the model is jointly statistically significant.
- If we are unable to reject the null hypothesis, we will say that the model is jointly statistically insignificant. In this case the average of y is a better estimator.
- Example:

$$\widehat{\log(price)} = \underset{(.570)}{.264} + \underset{(.151)}{1.043}\log(assess) + \underset{(.0386)}{.0074}\log(lotsize) - \underset{(.1384)}{.1032}\log(sqrft) + \underset{(.0221)}{.0338}\,bdrms$$

$$n = 88,\ \text{SSR} = 1.822,\ R^2 = .773.$$

- $$F = \frac{(R^2)/k}{(1-R^2)/(n-k-1)} = \frac{0.773/4}{(1-0.773)/83} = 70.66 > c$$

Testing General Linear Restrictions

- we can also test multiple restrictions without zeros.

$$H_0 : \beta_1 = 1, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$$

Testing General Linear Restrictions

- we can also test multiple restrictions without zeros.

$$H_0 : \beta_1 = 1, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$$

- Un-restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \mu$$

Testing General Linear Restrictions

- we can also test multiple restrictions without zeros.

$$H_0 : \beta_1 = 1, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$$

- Un-restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \mu$$

- Restricted Model:

$$(y - x_1) = \beta_0 + \mu$$

Testing General Linear Restrictions

- we can also test multiple restrictions without zeros.

$$H_0 : \beta_1 = 1, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$$

- Un-restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \mu$$

- Restricted Model:

$$(y - x_1) = \beta_0 + \mu$$

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim \mathbf{F}_{q,n-k-1}, \quad q = 4$$

Testing General Linear Restrictions

- we can also test multiple restrictions without zeros.

$$H_0 : \beta_1 = 1, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$$

- Un-restricted Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \mu$$

- Restricted Model:

$$(y - x_1) = \beta_0 + \mu$$

- $$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim \mathbf{F}_{q,n-k-1}, \quad q = 4$$

- The R-squared form cannot be used here, because the dependent variable in the restricted model $(y - x_1)$ differs from the one in the unrestricted model (y).

P-Value of F-Statistics

- We can also compute the P-value of F-Test following the same approach as we did in the t-test.

P-Value of F-Statistics

- We can also compute the P-value of F-Test following the same approach as we did in the t-test.
- $$P\text{-value} = P(\mathcal{F} > F)$$

  where $\mathcal{F}$ is a random variable following the $F_{q,n-k-1}$ distribution under H_0 and F is the observed F-statistic.

P-Value of F-Statistics

- We can also compute the P-value of F-Test following the same approach as we did in the t-test.
- $$P\text{-value} = P(\mathcal{F} > F)$$

  where $\mathcal{F}$ is a random variable following the $F_{q,n-k-1}$ distribution under H_0 and F is the observed F-statistic.
- The P-value is the smallest significance level at which you can reject the null hypothesis.

P-Value of F-Statistics

- We can also compute the P-value of F-Test following the same approach as we did in the t-test.
- $$P\text{-value} = P(\mathcal{F} > F)$$

  where $\mathcal{F}$ is a random variable following the $F_{q,n-k-1}$ distribution under H_0 and F is the observed F-statistic.
- The P-value is the smallest significance level at which you can reject the null hypothesis.
- **Rejection Rule**: $P - value < \alpha$