7. Multiple Regression Analysis: The Problem of Analysis

1Y & 2Xs

Three-Variable Model: Notation and Assumptions

Let us consider the following three-variable PRF as:

$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

where

Y_i is the dependent variable (regressand)

X_{2i} and X_{3i} are the regressors or the explanatory variables

u_i is the stochastic disturbance term

Remark: the subscript i is denoted the observation i from our sample data.

In case our data are time series, the subscript t will denote the t observation.

β_1 means the average value of Y when X_2 and X_3 are set equal to zero

β_2 and β_3 are called the partial regression coefficients.

We will talk about the meaning of β_2 and β_3 shortly after knowing the assumptions of the classical linear regression model (CLRM)

	Y_i	X_{2i}	X_{3i}
1980			
1981			
1982			
t	Y_i	X_{2i}	X_{3i}

Under the CLRM, we assume:

1. Zero mean value of u_i

$$E(u_i \mid X_{2i}, X_{3i}) = 0 \quad \equiv \text{Mean value of residuals} = 0$$

2. No serial correlation

$$Cov(u_i, u_j) = 0 \quad \text{for all } i \neq j$$

3. Homoscedasticity

$$Var(u_i) = \sigma^2$$

4. Zero covariance between u_i and each X variable, or

$$Cov(u_i, X_{2i}) = 0 \;;\; Cov(u_i, X_{3i}) = 0$$

5. No specification bias or

The model is correctly specified.

6. No exact collinearity between the X variables or

Specification bias comes from
1) Add redundant variable(s)
2) Omit important variable(s) in our model ("Exclusion bias")
3) Use wrong functional forms

Example of an exact linear relationship between Xs

X3i =2*X2i

Yi = B1 + B2X2i +B3X3i + ui

Yi = B1+ B2X2i + B3(2*X2i) +ui

Yi = B1 + (B2+2B3)X2i + ui

Observe that Partial Effect of X2i on Y and of X3i on Y cannot be estimated when X2i and X3i form an exact linear relationship

Exact linear relationship

Ex: X3i	X2i
2	1
4	2
6	3

X3i = 2*X2i

X3i - 2*X2i =0

Inexact linear relationship

X3i	X2i	X3i = 5*X2i + Vi
10	2	10 = 5*2 + 0
22	4	22 = 5*4 + 2
28	6	28 = 5*6 - 2
40	8	40 = 5*8 + 0

With Inexact linear relationship between X2i and X3i , OLS regression can be estimated, i.e., we can obtain B2_hat and B3_hat.

In general form: Lamda2*X2i + Lamda3*X3i =0

Here, in this example, Lamda2 = -2, Lamda3=1

Keypoint: If Lamda2 and Lamda3 are founded to make the above equation "TRUE" , then it is confirmed that X2 and X3 has "EXACT" linear relationship!
(We call this problem as " Perfect collinearity"

If this problem occurs, regression cannot be estimated.

By the above assumptions, we can find out the conditional expectation of Y_i:

Take $E[\cdot \mid X_{2i}, X_{3i}]$ to the PRF:

$$E[Y \mid X_{2i}, X_{3i}] = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + E[u_i \mid X_{2i}, X_{3i}]$$

(0)

The meaning of partial coefficients:

β_2

$$\frac{d\,E[Y \mid X_{2i}, X_{3i}]}{d X_{2i}} = \beta_2$$

holding X_3 constant, when X_2 changes by 1 unit, Y, on average, will change by β_2 unit.

β_3

$$\frac{d\,E[Y \mid X_{2i}, X_{3i}]}{d X_{3i}} = \beta_3$$

tell us partial effect of X_2 on Y

$$\frac{d\,E[Y|X_{2i},X_{3i}]}{dX_{2i}} = \beta_2$$

when X_2 changes by 1 unit, Y, on average, will change by β_2 unit. holding X_3 constant } partial effect of X_2 on Y

$$\frac{d\,E[Y|X_{2i},X_{3i}]}{dX_{3i}} = \beta_3$$

when X_3 changes by 1 unit, Y, on average, will change by β_3 unit. } partial effect of X_3 on Y

7.1 OLS Estimation of the Partial Regression Coefficients

In order to find the OLS estimators, we need to write down the sample regression function (SRF) corresponding to the PRF:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{u}_i$$

Goal: Find $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ that min $\sum \hat{u}_i^2$ (RSS $= \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$)

Trick: $\min_{\hat{\beta}_1,\hat{\beta}_2,\hat{\beta}_3} \sum \hat{u}_i^2 = \sum (Y_i - \hat{Y}_i)^2$

$$= \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \hat{\beta}_3 X_{3i})^2$$

FOCs:

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_1} = 0 \quad (1)$$

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_2} = 0 \quad (2)$$

$$\frac{\partial \sum \hat{u}_i^2}{\partial \hat{\beta}_3} = 0 \quad (3)$$

Three equations w/ three unknowns → we can solve for $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$

From the FOC, we then get the normal equations:

$$\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X}_2 + \hat{\beta}_3 \bar{X}_3$$
$$\sum Y_i X_{2i} = \hat{\beta}_1 \sum X_{2i} + \hat{\beta}_2 \sum X_{2i}^2 + \hat{\beta}_3 \sum X_{2i} X_{3i}$$
$$\sum Y_i X_{3i} = \hat{\beta}_1 \sum X_{3i} + \hat{\beta}_2 \sum X_{2i} X_{3i} + \hat{\beta}_3 \sum X_{3i}^2$$

We therefore get:

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}_2 - \hat{\beta}_3 \bar{X}_3$$
$$\hat{\beta}_2 = \frac{(\sum y_i x_{2i})(\sum x_{3i}^2) - (\sum y_i x_{3i})(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2}$$
$$\hat{\beta}_3 = \frac{(\sum y_i x_{3i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2}$$

Variance and Standard Errors of OLS Estimators

$$\text{var}(\hat{\beta}_1) = \left[\frac{1}{n} + \frac{\bar{X}_2^2 \sum x_{3i}^2 + \bar{X}_3^2 \sum x_{2i}^2 - 2\bar{X}_2\bar{X}_3 \sum x_{2i}x_{3i}}{\sum x_{2i}^2 \sum x_{3i}^2 - (\sum x_{2i}x_{3i})^2}\right] \sigma^2$$

$$\text{se}(\hat{\beta}_1) = +\sqrt{\text{var}(\hat{\beta}_1)}$$

$$\text{var}(\hat{\beta}_2) = \frac{\sum x_{3i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i}x_{3i})^2} \sigma^2$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2 (1 - r_{23}^2)}$$

$$\text{se}(\hat{\beta}_2) = +\sqrt{\text{var}(\hat{\beta}_2)}$$

$$\text{var}(\hat{\beta}_3) = \frac{\sum x_{2i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i}x_{3i})^2} \sigma^2$$

$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum x_{3i}^2 (1 - r_{23}^2)}$$

$$\text{se}(\hat{\beta}_3) = +\sqrt{\text{var}(\hat{\beta}_3)}$$

$$t = \frac{\hat{\beta}}{\text{se}(\hat{\beta})}$$

partial correlation coefficient bet. X_2 & X_3

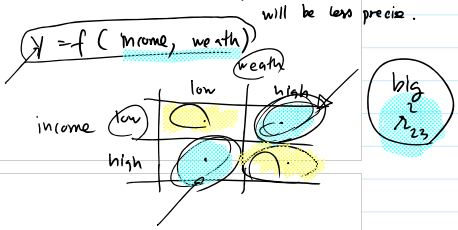
if r_{23}^2 is big, $\text{var}(\hat{\beta}_2)$ & $\text{var}(\hat{\beta}_3)$ will be large, ∴ our estimation will be less precise.

UNBIASNESS

$E(\hat{\beta}_1) = \beta_1$

$E(\hat{\beta}_2) = \beta_2$

$E(\hat{\beta}_3) = \beta_3$



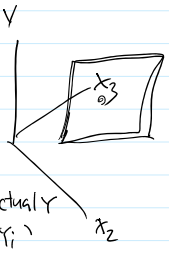
$\hat{\sigma}_u^2 \rightarrow \text{var}(\hat{\beta}_i) \rightarrow \text{se}(\hat{\beta}_i) \rightarrow t_{\hat{\beta}_i} \rightarrow$ hypothesis testing

$$\text{var}(\hat{\beta}_2, \hat{\beta}_3) = \frac{-r_{23}\sigma^2}{(1 - r_{23}^2)\sqrt{\sum x_{2i}^2}\sqrt{\sum x_{3i}^2}}$$

variance of our residuals: $\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-3} = \frac{RSS}{n-3}$ ($\hat{\sigma}_u^2$ (full description))

7.2 Properties of OLS Estimators: $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ — # of Xs including intercept term

① The three-variable regression model, the three-variable regression plane passes through the means $\bar{Y}, \bar{X}_2, \bar{X}_3$



② The mean value of the estimated Y_i ($\hat{Y}_i$) is equal to the mean value of actual Y (Y_i) (See the proof in the textbook)

$$\bar{\hat{Y}}_i = \bar{Y}_i$$

③ with deviation form: $y_i = \hat{y}_i + \hat{u}_i$

$$y_i = \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{u}_i$$

$$\hat{u}_i = y_i - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i}$$

$y_i = \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$

$\hat{u}_i = y_i - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i}$

$\sum \hat{u}_i^2 = \sum \hat{u}_i (y_i - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i})$

$\sum \hat{u}_i^2 = \sum \hat{u}_i y_i - \hat{\beta}_2 \sum \hat{u}_i x_{2i} - \hat{\beta}_3 \sum \hat{u}_i x_{3i}$ ($\sum \hat{u}_i x_{2i} = 0$, $\sum \hat{u}_i x_{3i} = 0$)

$= \sum y_i (y_i - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i})$

$\sum \hat{u}_i^2 = \sum y_i^2 - \hat{\beta}_2 \sum x_{2i} y_i - \hat{\beta}_3 \sum x_{3i} y_i$ → RSS

Properties of OLS Estimators (Cont.)

④ $\sum \hat{u}_i = 0$

and $\frac{\sum \hat{u}_i}{n} = \bar{\hat{u}}_i = 0$

⑤ $\hat{u}_i$ are uncorrelated with x_{2i} : $\sum \hat{u}_i x_{2i} = 0$

$\hat{u}_i$ " " x_{3i} : $\sum \hat{u}_i x_{3i} = 0$

⑥ $\hat{u}_i$ are uncorrelated with $\hat{Y}_i$: $\sum \hat{u}_i \hat{Y}_i = 0$

$x_i = X_i - \bar{X}$

$y_i = Y_i - \bar{Y}$

⑦ Recall that

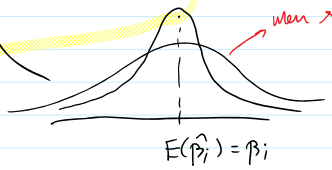
$$\text{var}(\hat{\beta}_2) = \frac{1}{\sum x_{2i}^2} \cdot \frac{\sigma^2}{(1 - r_{23}^2)}$$

$$\text{var}(\hat{\beta}_3) = \frac{1}{\sum x_{3i}^2} \cdot \frac{\sigma^2}{(1 - r_{23}^2)}$$

Given $\sum x_{2i}^2$ and σ^2, if r_{23}^2 ↑ → var($\hat{\beta}_2$) ↑ → Reduce Accuracy of prediction

If $r_{23}^2 = 1$, var($\hat{\beta}_2$) will be infinite [perfect multicollinearity]

In this case we cannot estimated $\hat{\beta}_2$



→ var($\hat{\beta}_i$) is high too

∴ $\hat{\beta}_2$ and $\hat{\beta}_3$ are "BLUE"

7.2 Properties of OLS Estimators

Properties of OLS Estimators (Cont.)

The Multiple Coefficient of Determination R^2 and the Multiple Coefficient of Correlation R

In this section, we will study how to measure the proportion of the variation in Y explained by the variables X_2 and X_3 jointly. This is the same concept of r^2 that we have learned before.

The quantity that gives this information is known as the **the multiple coefficient of determination** and is denoted by R^2.

To derive R^2, we firstly write down the following equation:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{u}_i = \hat{Y}_i + \hat{u}_i \quad (7.1)$$

→ STANDARD FORM OF OUR SRF

where $\hat{Y}_i$ is the estimated value of Y_i from the fitted regression line and is an estimator of true $E(Y_i|X_{2i}, X_{3i})$.

7.1 may be written as

$$y_i = \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{u}_i = \hat{y}_i + \hat{u}_i \quad (7.2)$$

→ DEVIATION FORM OF OUR SRF

Squaring 7.2 on both sides and summing over the sample values, we obtain

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum \hat{u}_i^2 + 2\sum \hat{y}_i \hat{u}_i = \sum \hat{y}_i^2 + \sum \hat{u}_i^2 \quad (7.3)$$

(since $\sum \hat{y}_i \hat{u}_i = 0$)

TSS ESS RSS

Recall that TSS = ESS + RSS → variation in Y_i that is left unexplained.

- TSS: Baseline or Total errors when using simple mean to estimate Y_i
- ESS: variation in Y_i that could be explained by the helps of X_{2i} and X_{3i}

$$TSS = ESS + RSS$$

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

$$R^2 = \text{Proportional variation in } Y \text{ that can be explained by the regression (i.e., by using } X_2 \text{ and } X_3)$$

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum \hat{u}_i^2}{\sum y_i^2} = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2}$$

Residual sum of squares (RSS); Total sum of squares (TSS)

Recall that $\sum \hat{u}_i^2 = \sum y_i^2 - \hat{\beta}_2 \sum x_{2i} y_i - \hat{\beta}_3 \sum x_{3i} y_i$

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum \hat{u}_i^2$$

(TSS = ESS + RSS)

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum y_i^2 - \hat{\beta}_2 \sum x_{2i} y_i - \hat{\beta}_3 \sum x_{3i} y_i$$

$$\sum \hat{y}_i^2 = \hat{\beta}_2 \sum x_{2i} y_i + \hat{\beta}_3 \sum x_{3i} y_i$$

→ ESS

$$R^2 = \frac{ESS}{TSS} = \frac{\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i}}{\sum y_i^2} \quad (7.4)$$

"R^2 in deviation form"

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \dots + \hat{\beta}_k X_{ki}$$

Regressors = $k-1$ variables

Define → R_j^2 as the R^2 in the regression of X_j on the remaining regressors ($k-1-1 = k-2$)

The three-or-more-variable analogue of r is the coefficient of multiple correlation, denoted by R, and it is a measure of the degree of association between Y and all the explanatory variables jointly. Although r can be positive or negative, R is always taken to be positive.

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{\sum x_j^2}\left(\frac{1}{1 - R_j^2}\right)$$

Ex: $R_{x_2}^2$ → $X_2 = \alpha_1 + \alpha_2 X_3$; $R_{x_3}^2$ → $X_3 = \alpha_1 + \alpha_2 X_2$

$\uparrow R_j^2 \rightarrow \uparrow Var(\hat{\beta}_j) \rightarrow \uparrow s.e.(\hat{\beta}_j) \rightarrow \hat{t} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$ (low $\hat{t}$)

$H_0: \beta_j = 0$

$H_a: \beta_j \neq 0$

7.2.1 R^2 and the Adjusted R^2

It should be noted that the R^2 is a nondecreasing function of the number of explanatory variables. Thus, when the number of regressors increases, R^2 almost invariably increases and never decreases. **In other words, an additional X variable will not decrease R^2.**

To explain this fact, let we write down the definition of R^2 again:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum \hat{u}_i^2}{\sum y_i^2} \quad (7.5)$$

Therefore, in comparing two regression models **with the same dependent variable but differing number of X variables**, one should be very wary of choosing the model with the highest R^2.

In light of comparing two R^2 terms, we have to take into account the number of X variables present in the model. To achieve this goal , we can consider the alternative coefficient of determination, which is as follows:

Adjusted R^2: $\bar{R}^2 = 1 - \frac{\sum \hat{u}_i^2 / (n-k)}{\sum y_i^2 / (n-1)}$ (for $k = 3$)

In light of comparing two R^2 terms, we have to take into account the number of X variables present in the model. To achieve this goal , we can consider the alternative coefficient of determination, which is as follows:

Adjusted R^2 :

$$\overline{R^2} = 1 - \frac{\sum \hat{u}_i^2 / (n-k)}{\sum y_i^2 / (n-1)}$$

for $k = 3$

General form :

$$\overline{R^2} = 1 - \frac{\sum \hat{u}_i^2 / (n-k)}{\sum y_i^2 / (n-1)}$$

k= the number of parameters in the model including the intercept term.
n= the number of observations in the sample data.

The above equation is known as **the adjusted** R^2, denoted by $\bar{R}^2$. The term adjusted means adjusted for the df associated with the sums of squares entering into 7.5.

We can rewrite the the adjusted R^2 as:

$$\overline{R^2} = 1 - \frac{\hat{\sigma}^2}{S_Y^2} \qquad \text{where} \quad \hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-k} \rightarrow \text{unbiased estimator of } \sigma^2$$

$$S_Y^2 = \frac{\sum y_i^2}{n-1} \rightarrow \text{sample variance of } Y$$

We can also get the equation which shows the relationship between R^2 and $\bar{R}^2$

$$\overline{R^2} = 1 - \frac{\sum \hat{u}_i^2}{\sum y_i^2} \cdot \frac{(n-1)}{(n-k)}$$

$$\overline{R^2} = 1 - [1 - R^2] \cdot \frac{(n-1)}{(n-k)} \rightarrow \text{As } k > 1 \rightarrow \frac{(n-1)}{(n-k)} > 1$$

NOTE $\overline{R^2}$ cannot be negative.

$\frac{(n-1)}{(n-5)}$ $\quad \frac{(n-1)}{(n-3)} = 1$

$\frac{10-1}{10-3} = \frac{9}{6} > 1$

Besides R^2 and $\bar{R}^2$ as goodness of fit measures, other criteria are often used to judge the adequacy of a regression model. Two of these are **Akaike's Information criterion and Amemiya's Prediction criteria**, which are used to select between competing models. We will discuss these criteria in greater detail later.

Result: $\overline{R^2} < R^2$ ***

Remarks When comparing two R^2 values, make sure that

① Dependent variable of the two models is the same

② Sample size must be the same

Ex:

$$\ln Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} \quad (1)$$

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} \quad (2)$$

→ ① and ② cannot be directly compared by using R^2 b/c the dependent variables are NOT the same!

8. Multiple Regression Analysis: The Problem of Inference

In this chapter, we will extend the ideas of interval estimation and hypothesis testing developed there to models involving three or more variables.

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

We have already known that if our objective is to do interval estimation and hypothesis testing, we need to assume that the u_i follow the normal distribution with zero mean and constant variance σ^2.

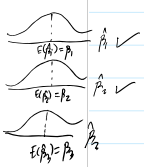
With the normality assumption and the CLRM assumptions, we know that:

[1] The OLS estimations of partial regression coefficients are best linear unbiased estimators (BLUE).

[2] The estimators $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ are normally distributed with means equal to true β_1, β_2, and β_3 and variances are following:

$$var(\hat{\beta}_1) = \left[\frac{1}{n} + \frac{\bar{X}_2^2 \sum x_{3i}^2 + \bar{X}_3^2 \sum x_{2i}^2 - 2\bar{X}_2\bar{X}_3 \sum x_{2i}x_{3i}}{\sum x_{2i}^2 \sum x_{3i}^2 - (\sum x_{2i}x_{3i})^2}\right] \cdot \sigma^2$$

$$se(\hat{\beta}_1) = +\sqrt{var(\hat{\beta}_1)}$$



$$var(\hat{\beta}_2) = \frac{\sum \hat{r}_{i1}^2}{(\sum \hat{r}_{i1}^2)(\sum \hat{r}_{i2}^2) - (\sum \hat{r}_{i1}\hat{r}_{i2})^2} \cdot \sigma^2$$

$$var(\hat{\beta}_2) = \frac{\sigma^2}{\sum \hat{r}_{i1}^2 (1 - r_{12}^2)}$$

$$se(\hat{\beta}_2) = +\sqrt{var(\hat{\beta}_2)}$$

$$var(\hat{\beta}_1) = \frac{\sum \hat{r}_{i2}^2}{(\sum \hat{r}_{i1}^2)(\sum \hat{r}_{i2}^2) - (\sum \hat{r}_{i1}\hat{r}_{i2})^2} \cdot \sigma^2$$

$$var(\hat{\beta}_1) = \frac{\sigma^2}{\sum \hat{r}_{i2}^2 (1 - r_{12}^2)}$$

$$se(\hat{\beta}_1) = +\sqrt{var(\hat{\beta}_1)}$$

Moreover, $\frac{(n-k)\hat{\sigma}^2}{\sigma^2}$ follows the χ^2 distribution with n-3 df. We can also show that, if we replace the true σ^2 by its unbiased estimator $\hat{\sigma}^2$ in the computation of the standard errors, we then get

$$t = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)}$$
$$t = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)}$$
$$t = \frac{\hat{\beta}_3 - \beta_3}{se(\hat{\beta}_3)}$$

follows the t distribution with n-3 df. $n-k = n-3$

Example Consider the following regression:

$$\underset{se}{\log(\widehat{salary})} = \underset{(0.32)}{4.32} + \underset{(0.035)}{0.280}\log(sales) + \underset{(0.0041)}{0.0174}\,ROE + \underset{(0.00054)}{0.00024}\,ROS \quad (8.1)$$

$$R^2 = 0.283$$

where
salary =salary of CEO (thousand dollars)
sales=annual firm sales
ROE=return on equity in percent
ROS =return on firm's stock

Interpreter the partial regression coefficients

Questions What about the statistical significance of the observed results?

For the coefficient of log(sales) of 0.280, Is this coefficient statistically significant different from zero?

For the coefficient of ROE of 0.0174, Is this coefficient statistically significant different from zero?

For the coefficient of ROS of 0.00024, Is this coefficient statistically significant different from zero?

Are these three coefficients statistically significant?
To answer these questions, we have to learn the kinds of hypothesis testing.

8.1 Hypothesis Testing About Individual Regression Coefficients

We can use the t-test to test a hypothesis about any individual partial regression coefficient.

8.1.1 Two-tail test:

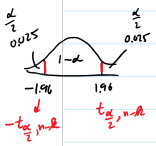
Let us postulate that
H_0: (sales do not affect CEO salary)
H_1: $\beta_2 \neq 0$ (sales do affect CEO salary)

hypothesized value in H_0

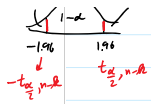
suppose $\alpha = 0.05$

① Compute $\hat{t}$: $\hat{t} = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \frac{0.28 - 0}{0.035} = 8.00$

② find critical t : $n = 209$ firms
$df = n-k = 209 - 4 = 205$
$t_{\frac{\alpha}{2}, n-k} = t_{\frac{0.05}{2}, 205} = 1.96$



$t_{\frac{\alpha}{2}, n-k} = t_{\frac{0.05}{2}, 205} = 1.96$



③ Compare $\hat{t}$ w/ critical t: Since $\hat{t}$ > critical t, then we can reject the null hypothesis that $H_0: \beta_2 = 0$ in favor of the alternative hypothesis that $H_a: \beta_2 \neq 0$.

In other words, $\hat{\beta}_2$ is statistically significant different from zero at 95%.

130 Chapter 8. Multiple Regression Analysis: The Problem of Inference

8.1 Hypothesis Testing About Individual Regression Coefficients 131

8.1.2 One-tail tests:

Let us postulate that

$H_0: \beta_2 \leq 0$
$H_1: \beta_2 > 0$

(sales do have a positive impact on CEO salary)

$t_{\alpha, df} = t_{0.05, 205} = 1.645$

[D - I - Y]

132 Chapter 8. Multiple Regression Analysis: The Problem of Inference

8.2 Testing The Overall Significance of the Sample Regression [Joint Test]

In the previous section, we test the significance of the estimated partial regression coefficients individually, that is under the separate hypothesis that each true population partial regression coefficient was zero. But now we are interested in testing β_2 , β_3 and β_k are jointly or simultaneously equal to zero. In other words, we would like to test the following hypothesis:

$$H_0: \beta_2 = \beta_3 = \beta_k = 0$$

H_a : At least one of our explanatory variables is NOT equal to zero.

In order to reach this goal, we have to learn the following test.

The Analysis of Variance Approach to Testing the Overall Significance of an Observed Multiple Regression: The F-Test

The joint hypothesis can be tested by the **Analysis of Variance (ANOVA)** which can be demonstrated as follows:

Table 8.1: ANOVA Table for the three-variable regression model

Source of variation	Sum of Square SS	df	Mean Sum of Square MSS
① Due to regression (ESS)	$\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i}$	2	$(\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i})/2$
② Due to residuals (RSS)	$\sum \hat{u}_i^2$	$n-3$	$\sum \hat{u}_i^2/(n-3)$
① + ② = ③ TSS	$\sum y_i^2$	$n-1$	$\sum y_i^2/(n-1)$

① = (k-1)

② = ① / ②

$$TSS = ESS + RSS$$
$$= \hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i} + \sum \hat{u}_i^2$$

For Joint test

Let's say $H_0: \beta_2 = \beta_3 = 0$ Both X_2 and X_3 could not jointly explain variation in Y

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

Let's compute F- statistic :

$$F = \frac{ESS/df}{RSS/df} = \frac{(\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i})/2}{\sum \hat{u}_i^2/(n-3)}$$

$k-1 = 3-1$

$=k$

b/c, to compute $\sum \hat{u}_i^2$, we have to estimate $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ which consume 3 degrees of freedom

So, $df = n-3$.

8.2 Testing The Overall Significance of the Sample Regression 133

Decision Rule Given the k- variable regression model:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i$$

To test the hypothesis

$$H_0: \beta_2 = \beta_3 = \ldots = \beta_k = 0$$

(i.e ., all slope coefficients are simultaneously zero) versus

H_1: Not all slope coefficients are simultaneously zero

If $F > F_\alpha(k-1, n-k)$, we reject H_0 ; otherwise we cannot reject it, where $F_\alpha(k-1, n-k)$ is the critical F value at the α level of significance and (k-1) numerator df and (n-k) denominator df.

An important Relationship between R^2 and F

$TSS = ESS + RSS$

$R^2 = \frac{ESS}{TSS}$

$$F = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{ESS}{RSS}\cdot\frac{(n-k)}{(k-1)} = \frac{ESS}{TSS-ESS}\cdot\frac{(n-k)}{(k-1)} = \frac{\frac{ESS}{TSS}}{\frac{TSS-ESS}{TSS}}\cdot\frac{(n-k)}{(k-1)}$$

$$F = \frac{R^2}{1-R^2}\cdot\frac{(n-k)}{(k-1)}$$

$$\boxed{F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}}$$

① If $R^2 = 0$, then $F = 0$

② The higher the R^2, The greater the F-value.

③ In limit when $R^2 \to 1$, F becomes infinite.

k = # of explanatory variable including intercept.

Ex: $H_0: \beta_2 = \beta_3 = \beta_4 = 0$

H_a: Not all slope coefficients are zero simultaneously

(sales, ROE, ROS)

STEP 1: Compute F-statistic:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.28/(4-1)}{(1-0.28)(209-4)} = 26.9712$$

STEP 2: Find critical F stat:

$\alpha = 0.05 \rightarrow F_{\alpha, k-1, n-k} = F_{005, 4-1, 209-4} = F_{0.05, 3, 205} = 2.65$

STEP 3: Compare $\hat{F}$ w/ critical F-stat:

Since $\hat{F} = 26.9712 > F_{CRITICAL} = 2.65$, then we may comfortable reject $H_0: \beta_2 = \beta_2 = \beta_4 = 0$ in favor of H_a at 95% confidence level.

8.2 Testing The Overall Significance of the Sample Regression 135

Table 8.3 ANOVA Table in Terms of R^2 $(k-1) = (3-1)$

Source of variation	Sum of Square (SS)	df	Mean Sum of Square (MSS)
Due to regression (ESS)	$R^2 \cdot \sum y_i^2$	2	$(R^2 \cdot \sum y_i^2)/2$
Due to residuals (RSS)	$(1-R^2)\cdot \sum y_i^2$	$n-3$	$[(1-R^2)\cdot \sum y_i^2]/(n-3)$
TSS	$\sum y_i^2$	$n-1$	$\sum y_i^2/(n-1)$

Decision Rule Testing the overall significance of a regression in terms of R^2

Given the k- variable regression model:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i$$

To test the hypothesis

H_0: $\beta_2 = \beta_3 = \ldots = \beta_k = 0$

(i.e .. all slope coefficients are simultaneously zero) versus

H_1: Not all slope coefficients are simultaneously zero

Compute

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

If $F > F_\alpha(k-1, n-k)$, we reject H_0; otherwise we cannot reject it, where $F_\alpha(k-1, n-k)$ is the critical F value at the α level of significance and (k-1) numerator df and (n-k) denominator df.