Model 2:
\[ \hat{QD} = 1.52 \text{ (Income)} - 6.53 \quad R^2 = 0.612 \]

Model 3:
\[ \hat{QD} = -2.105 - 0.579 \text{ Price} + 4.075 \text{ Income} \quad R^2 = 0.841 \]

\[ \text{Var}(\hat{\beta}_0) = 5.61 \]
\[ \text{Var}(\hat{\beta}_1) = 0.011 \]
\[ \text{Var}(\hat{\beta}_2) = 0.243 \]
\[ \text{Var}(\hat{\epsilon}) = \sigma^2 = (1.97)^2 \]

SLR of Price on Income
\[ \hat{I} = 10.0 + 0.2 \text{ Price} \quad R^2 = (0.939)^2 \]

This shows two independent variables in model 3 do have strong correlation.
This situation is called "Multicollinearity".

22/05/2012

Suppose Model I:
\[ \hat{QD} = \hat{\beta}_0 + \hat{\beta}_1 \text{ Price} = -2.11 + 0.24 \text{ Price} \]
Model II:
\[ \hat{QD} = \hat{\beta}_0 + \hat{\beta}_1 \text{ Income} = -6.53 + 1.52 \text{ Income} \]
Model III:
\[ \hat{QD} = -2.105 + 4.075 \text{ Income} - 0.579 \text{ Price} \]
\[ \hat{QD} = \hat{\beta}_0 + \hat{\beta}_1 \text{ Price} + \hat{\beta}_2 \text{ Income} \]

Model IV:
\[ \hat{\text{Income}} = \hat{\beta}_0'' + \hat{\beta}_1'' \text{ Price} \]
\[ = 10 + 0.2 \text{ Price} \]

Inserting Model IV into Model III:
\[ \hat{QD} = \hat{\beta}_0 + \hat{\beta}_1 \text{ Price} + \hat{\beta}_2 (\hat{\beta}_0'' + \hat{\beta}_1'' \text{ Price}) \]
\[ = \hat{\beta}_0' + \hat{\beta}_1' \text{ Price} + \hat{\beta}_2' \hat{\beta}_0'' + \hat{\beta}_2' \hat{\beta}_1'' \text{ Price} \]
\[ = (\hat{\beta}_0' + \hat{\beta}_2' \beta_0'') + (\hat{\beta}_1' + \hat{\beta}_2' \beta_1'') \text{ Price} = B_0 + B_1 \text{ Price} \]
Omitted Variable Bias

Suppose \( Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + U \) is the population model.

But we omit \( x_2 \) from the model and assume

\( \tilde{Y} = \beta_0 + \beta_1 x_1 + \nu \)

where \( \nu = \beta_2 x_2 + U \)

When we estimate the parameters of \( \tilde{Y} \),
\( \hat{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \tilde{s}_1 \)

where \( X_2 = \tilde{s}_1 x_1 \), \( \tilde{s}_1 \) including the coefficient of \( x_1 \) on \( x_2 \)

Taking expectations

\[
E[\hat{\beta}_1] = E[\beta_1 + \hat{\beta}_2 \tilde{s}_1]
\]

\[
= E[\hat{\beta}_1] + \tilde{s}_1 E[\hat{\beta}_2]
\]

\[
E[\hat{\beta}_1] = \beta_1 + \tilde{s}_1 \beta_2 , \text{ that is:}
\]

\[
E[\hat{\beta}_1] - \beta_1 = \tilde{s}_1 \beta_2 \quad \text{yielding a bias of } \tilde{s}_1 \beta_2
\]

Including Irrelevant Variables

\( Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + U \)

\( Y \): cell phone bill \( X_3 \): ice cream consumption

Suppose \( X_3 \) has no effect on \( Y \) when \( X_1, X_2 \) are held fixed

i.e.: \( \beta_3 = 0 \), then \( E[Y | x_1 x_2 x_3] = E[Y | x_1 x_2] \)

Question: How can we decide on the significance of the parameters and then the independent variables?

If we can justify that \( \beta_3 = 0 \), we can say \( X_3 \) has no impact on \( Y \).
CHAPTER 4
Multivariate Regression - Hypothesis Testing

Null hypothesis \[ H_0: \beta_k = 0 \quad k=0,1,\ldots \]

Alternative Hypoth. \[ H_A: \beta_k \neq 0 \]

Steps to justify the claim
1. Collect data / information
2. Estimate parameter, \( \hat{\beta}_k \)
3. Check if \( \hat{\beta}_k \) is in acceptable range of population
4. Calculate test statistics
5. Calculate p-value

Important Measure: Significance Level \( \alpha \)
\[ \alpha = \text{probability of Rej.} \quad \text{Ho when Ho is True} \]

\( (1-\alpha) \) is the confidence level that \( \beta_k \) will be within \( A \) and \( B \).

\( A \) and \( B \) are functions of \( \hat{\beta}_k \). And how to determine their values?

Symmetrical Distribution

1. Normal Distribution \( X \sim N(\mu, \sigma^2) \)
Standard Normal Distribution \[ Z \sim N(0,1) \]

Student's t-Distribution \( t \sim t_{df, \alpha} \) \( df \): degree of freedom

Standardized Value of a random variable
\[
z = \frac{X - E[X]}{S.\text{Dev.}[X]} = \frac{X - \mu}{\sigma}
\]

Standardized Value of a parameter
\[
z = \frac{\hat{\beta}_k - E[\hat{\beta}_k]}{\text{st. Dev.}(\hat{\beta}_k)} = \frac{\hat{\beta}_k - \beta_k}{\sqrt{\text{Var}(\hat{\beta}_k)}}
\]

For small \( n \), unknown variance \( \Rightarrow t\)-distribution
For large \( n \) \( \Rightarrow z\)-distribution

**TEST STATISTICS**

We claim \( H_0: \hat{\beta}_k = \beta_{ko} \) \( \beta_{ko} \): a value claimed for \( \beta_k \)

\( H_A: \hat{\beta}_k \neq \beta_{ko} \)
Test Statistic = \( \frac{\hat{\beta}_k - \beta_{k0}}{\sqrt{\text{Var}(\hat{\beta}_k)}} \)

when \( H_0: \beta_k = 0 \) then \( \beta_{k0} = 0 \), then

Test Statistic = \( \frac{\hat{\beta}_k}{\sqrt{\text{Var}(\hat{\beta}_k)}} \)

Example: QD, Price, Income example from Handout 2

Model 3: \( QD = \beta_0 + \beta_1 P + \beta_2 I + u \)

\( \hat{\beta}_0 = -2.105 \quad \hat{\beta}_1 = -0.57 \quad \hat{\beta}_2 = 4.075 \)

\( \text{Var} \hat{\beta}_0 = 2.37 \quad \text{Var} \hat{\beta}_1 = 0.104 \quad \text{Var} \hat{\beta}_2 = 0.493 \)

\( H_0: \beta_0 = 0 \)

\( H_A: \beta_0 \neq 0 \)

Test statistics \( t = \frac{-2.105}{2.37} = -0.8846 \)

Rejection Region for a given significance level \( \alpha = 0.05 \)

\( n = 25 \) obs.

\( t_{24, 0.025} = 2.0609 \)

From \( t \)-table

\(-0.8846 \) is in the acceptance region. Therefore, we fail to reject \( H_0: \beta_0 = 0 \). This means intercept \( \beta_0 \) is not significant in the model.

Given the claim \( H_0: \beta_1 = 0 \) vs \( H_A: \beta_1 \neq 0 \)

Test statistics \( t = \frac{-0.57}{0.104} = -5.51 \)

\(-5.51 \) lies within rejection region \( \Rightarrow \) reject \( H_0 \). That is \( \beta_1 \), therefore, Price, is significant in the model.
Multiple Linear Regression

We set assumptions on OLS estimators

1. Linearity in parameters
2. Random sample of n-observation
3. No perfect Collinearity
4. Zero Conditional Mean $E[u|x] = 0$

Assumptions 1-4 guarantees unbiasedness of $\hat{\beta} = [\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_j, \ldots, \hat{\beta}_k]$

$E[I\hat{\beta}_j] = \beta_j$

5. Homoskedasticity

$\text{Var}[u|x] = \sigma^2$

Assumptions 1-5 Gauss Markov Theorem:

$\hat{\beta}$ is BLUE and minimizes the variance of random error.

Which other assumptions we need to justify the model?

1. Given $y$, population model, what is the distribution of $u$ and $\hat{u}$

$u \sim N(0, \sigma^2)$ Normally distributed with mean $\mu$, variance $\sigma^2$

The asymptotic distribution of $\hat{u} \sim N(0, \sigma^2)$

2. Given $\hat{\beta}$'s are unbiased & BLUE, what is the sampling distribution of $\hat{\beta}_j$?

$\hat{\beta}_j \overset{\text{Asym.}}{\sim} N(\beta_j, \text{Var}[\hat{\beta}_j])$

$\hat{\beta}_j$ Asymptotically Normal with mean $\beta_j$, variance $\text{Var}[\hat{\beta}_j]$
Why we use t-distribution:

1. No perfect information on population variance $\sigma^2$. We estimate it by using $\text{Var}(\hat{U}) = \hat{\sigma}^2$.

   t-distribution uses the information on the dist. of $\hat{\sigma}^2$.

2. It represents small sample data much better than other symmetrical distributions.

3. For large $n$, it converges to normal distribution.

Under Assumptions 1-5, we can assume

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\text{Var}(\hat{\beta}_j)}} \sim t_v$$

where $\hat{\beta}_j$ is the sampling distribution of the parameter estimate $eta_j$.

How to express $\hat{\beta}_j$ and $\text{Var}(\hat{\beta}_j)$ in terms of its components?

$$\hat{\beta}_j = \frac{\sum \hat{r}_{ij} y_i}{\sum \hat{r}_{ij}^2}$$

Partially Out

where $\hat{r}_{ij}$ : residuals from regression of $X_j$ on other independent variables.

Example: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

$$\hat{\beta}_1 = \frac{\sum \hat{r}_{1i} y_i}{\sum \hat{r}_{1i}^2}$$

How to find $\hat{r}_{1i}$'s?
Regress $X_1$ on $X_2$:

$X_1 = \beta_0 + \beta_1 X_2 + \epsilon_1$

$\hat{X}_1 = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}_2$  

$\Rightarrow \hat{\epsilon}_1 = X_1 - \hat{X}_1$

$$\text{Var} [\hat{\beta}_j] = \sigma^2 \left[ \frac{1}{\sum (x_{ij} - \bar{x}_j)^2 (1 - R_i^2)} \right]$$  

$= \frac{\sigma^2}{\text{SST}_y (1 - R_i^2)}$

$R_i^2$: Coefficient of Determination for $X_i$ on all other $X_j$

Example:

$$\text{Var} [\hat{\beta}_1] = \frac{\sigma^2}{\sum (x_{1j} - \bar{x}_1)^2 (1 - R_i^2)}$$

higher the $R_i^2$, higher $\text{Var} [\hat{\beta}_1]$

Back to hypothesis Testing

$H_0: \beta_j = 0$  

$H_A: \beta_j > 0$  

$\downarrow$  

TWO-SIDED

$H_0: \beta_j = 0$  

$H_A: \beta_j \\Leftrightarrow 0$  

$\downarrow$  

ONE-SIDED
Steps are

1. Set the hypothesis  
   \( H_0: \beta_j = 0 \)  
   \( H_A: \beta_j \neq 0 \)  
   \( j = 0, 1, 2, \ldots, k \)

2. Collect data and estimate \( \hat{\beta}_j \)

3. Calculate test statistics  
   \[ t = \frac{\hat{\beta}_j - 0}{\text{s.e.}(\hat{\beta}_j)} \]

4. Set a significance level \( \alpha \) and find value(s) for critical region \( c \)

5. If \( |t| > c \) reject \( H_0 \Rightarrow \beta_j \neq 0 \)  \( \beta_j \) is significant  
   If \( |t| < c \) fail to reject \( H_0 \Rightarrow \beta_j \) is not significant in the equation.

Example: Handout 3 p: 6

AD vs P, I

1. \( H_0 : \beta_2 = 0 \) vs \( H_A : \beta_2 \neq 0 \)
2. \( \hat{\beta}_2 = 4.075 \)
3. \( t = \frac{4.075}{0.49} = 8.26 \)
4. Set \( \alpha = 5\% \) \( n = 24 \)  
   \[ t_{23, 0.025} = 2.069 \]  
   \( c = 2.069 \) critical value

5. \( 8.26 > 2.069 \) reject \( H_0 \).  
   Income is significant in the model.
One-sided test

1. \( \text{Ho} : \beta_2 = 0 \)
   \( \text{Ha} : \beta_2 > 0 \)

2. \( \hat{\beta}_2 = 4.075 \)

3. \( t = 8.26 \)

4. \( \alpha = 5\% \quad C = t_{23,0.05} = 1.714 \)

5. \( 8.26 > 1.714 \) \text{Reject Ho.}

Definition: Given the observed value of the test statistic, what level of significance is achieved in its smallest value is called p-value.

\[ p \left( |t| > \text{test stat.} \right) = p\text{-value} \]

Example: Suppose test-stat. = 2.45 \( n = 30 \)
\[ p \left( |t| > 2.45 \right) = 0.02 \quad \text{in two-sided test} \]
\[ p \left( t > 2.45 \right) = 0.01 \quad \text{one-sided.} \]

Example: OD, P, I example
\( \text{Ho: } \beta_0 = 0 \)
\( \text{Ha: } \beta \neq 0 \)

\( p\text{-value} = 0.39 \geq \text{any } \alpha \Rightarrow \text{Fail to reject } \text{Ho.} \)

that is \( Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u \) can be taken as \( Y = \beta_1 X_1 + \beta_2 X_2 + u \) \( \Rightarrow \text{Regression through origin} \)
Regression through origin

\[ \hat{y} = \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k + u \quad \hat{\beta}_0 = 0 \]

\[ k=1 \Rightarrow \hat{y} = \hat{\beta}_1 x + u \]

and

\[ \hat{\beta}_1 = \frac{\sum \hat{y}_i x_i}{\sum x_i^2} \]

When \( k > 0 \): Having \( \hat{\beta}_0 = 0 \) leads to

1. \( \sum \hat{u}_i \neq 0 \)
2. \( R^2 = 1 - \frac{SSR}{SST} < 0 \)

That is, \( \hat{y} \) explains more variation than \( x_1 \ldots x_k \)

Confidence Intervals.

When we set \( \hat{\beta}_2 = 4.075 \), it gives a unique value for \( \beta_2 \) which on average equals to \( \beta_2 \).

\( (1-\alpha)100\% \) confidence interval on \( \beta_j \) is

\[ P(A \leq \hat{\beta}_j \leq B) = 1-\alpha \]

We know

\[ P(1 \leq \chi^2 \leq C) = \alpha \]
\[ P(-c \leq t \leq c) = 1-\alpha \]
\[ P(-c \leq \frac{\hat{\beta}_j - \beta_j}{\text{s.e.}\hat{\beta}_j} \leq c) = 1-\alpha \]
\[ P(\hat{\beta}_j - c \text{s.e.}\hat{\beta}_j \leq \beta_j \leq \hat{\beta}_j + c \text{s.e.}\hat{\beta}_j) = 1-\alpha \]
Example: Find 95% confidence interval for $\beta_2$, if

\[
\text{Sales} = 16.406 - 8.247 \text{Price} + 0.585 \text{Ads}
\]

\[
(4.34) \quad (2.196) \quad (0.1336) \implies \text{s.e.}
\]

\[
P \left( \frac{-1.96 \text{ s.e} \beta_2}{\beta_2} \leq \beta_2 \leq \frac{+1.96 \text{ s.e} \beta_2}{\beta_2} \right) = 0.95
\]

where $c=1.96$ when $(1-\alpha)$ is 95%.

\[
\hat{\beta}_2 \pm 1.96 \text{ s.e} (\hat{\beta}_2)
\]

\[
0.585 \pm 1.96 (0.1336)
\]

Prob. that $\beta_2$ will lie within

$0.585 - 1.96 (0.1336)$ and $0.585 + 1.96 (0.1336)$ is 0.95.

**Testing More than one parameter**

Besides single parameter tests, we can compare and test the equality of the parameters such as

$H_0: \beta_1 = \beta_2 \implies \beta_1 - \beta_2 = 0$

$H_0: \beta_2 = \beta_3 \implies \beta_2 - \beta_3 = 0$

$H_0: \beta_3 = \beta_1 \implies \beta_3 - \beta_1 = 0$

$\implies H_0: \beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$ \text{ Test on all parameters simultaneously}

$H_4: \text{At least one differs}$

Test statistic $= F$ with degrees of freedom 1 and d.o.f 2

$F_{1, 2}$

**Example:** $Q$, $p$, $I$ $F = 55.67$

$P \left( F > 55.67 \right) = 0 \implies \text{Reject } H_0: \beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$

The model is linear.