Chapter 5
The Multiple Regression Model

Walter R. Paczkowski
Rutgers University

Chapter Contents

5.1
Introduction

- Let's set up an economic model in which sales revenue depends on one or more explanatory variables
  - We initially hypothesize that sales revenue is linearly related to price and advertising expenditure
  - The economic model is:

Eq. 5.1

$$SALES = \beta_1 + \beta_2 PRICE + \beta_3 ADVERT$$

- In most economic models there are two or more explanatory variables
  - When we turn an economic model with more than one explanatory variable into its corresponding econometric model, we refer to it as a **multiple regression model**
  - Most of the results we developed for the simple regression model can be extended naturally to this general case

- β_2 is the change in monthly sales *SALES* ($1000) when the price index *PRICE* is increased by one unit ($1), and advertising expenditure *ADVERT* is held constant

$$\beta_2 = \frac{\Delta SALES}{\Delta PRICE\ (ADVERT \text{ held constant})}$$

$$= \frac{\partial SALES}{\partial PRICE}$$

- Similarly, β_3 is the change in monthly sales *SALES* ($1000) when the advertising expenditure is increased by one unit ($1000), and the price index *PRICE* is held constant

$$\beta_3 = \frac{\Delta SALES}{\Delta ADVERT \left(PRICE \text{ held constant} \right)}$$

$$= \frac{\partial SALES}{\partial ADVERT}$$

- The economic model in Eq. 5.1 should be written as:

$$E(SALES) = \beta_1 + \beta_2 PRICE + \beta_3 ADVERT$$

  - To allow for a difference between observable sales revenue and the expected value of sales revenue, we add a random error term, $e = SALES - E(SALES)$

Eq. 5.2

$$SALES = E(SALES) + e = \beta_1 + \beta_2 PRICE + \beta_3 ADVERT + e$$

FIGURE 5.1 The multiple regression plane

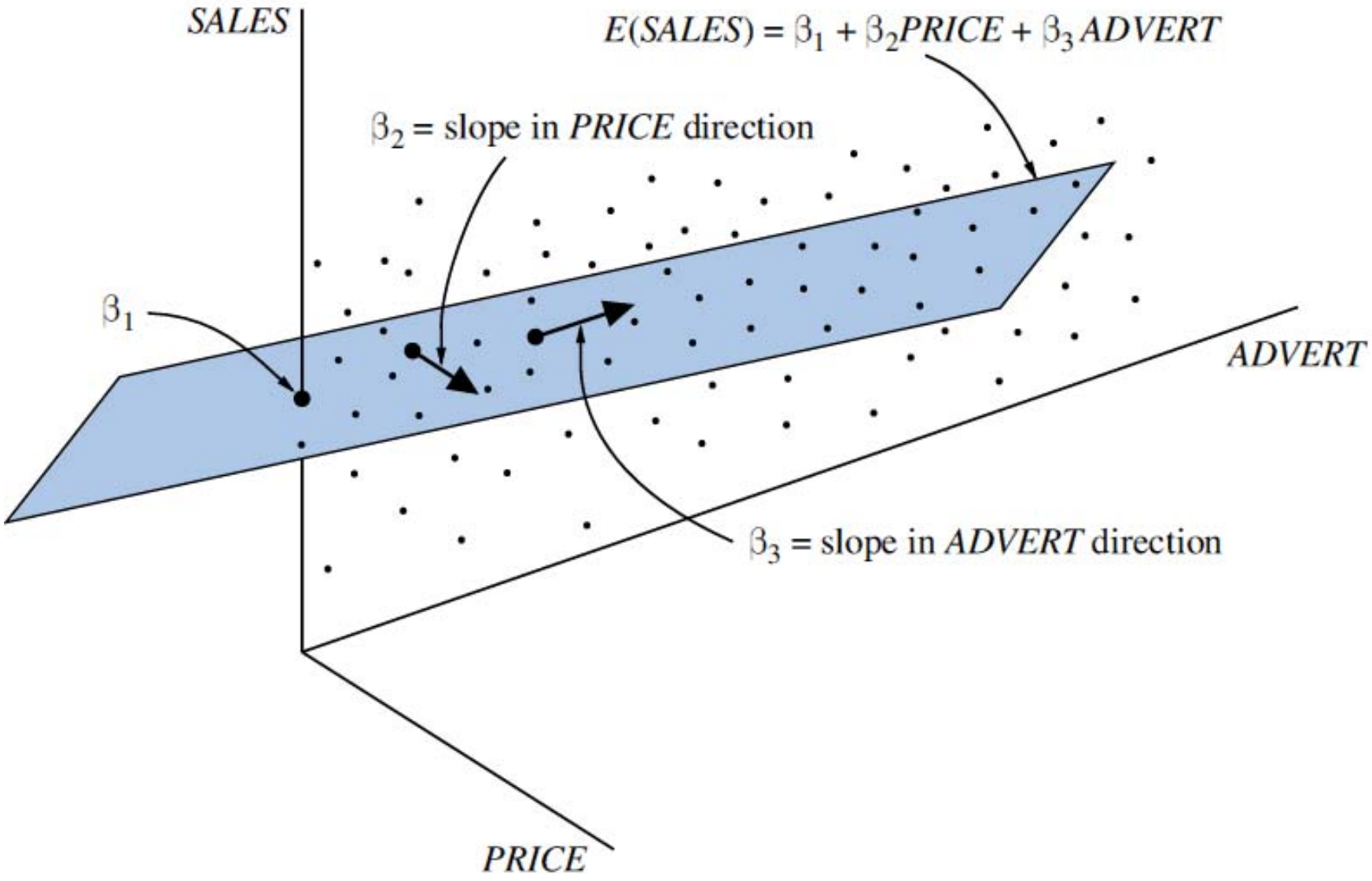

Table 5.1 Observations on Monthly Sales, Price, and Advertising in Big Andy's Burger Barn

City	*SALES* $1,000 units	*PRICE* $1 units	*ADVERT* $1,000 units
1	73.2	5.69	1.3
2	71.8	6.49	2.9
3	62.4	5.63	0.8
4	67.4	6.22	0.7
5	89.3	5.02	1.5
.	.	.	.
.	.	.	.
.	.	.	.
73	75.4	5.71	0.7
74	81.3	5.45	2.0
75	75.0	6.05	2.2
	Summary statistics		
Sample mean	77.37	5.69	1.84
Median	76.50	5.69	1.80
Maximum	91.20	6.49	3.10
Minimum	62.40	4.83	0.50
Std. Dev.	6.49	0.52	0.83

- In a general multiple regression model, a dependent variable y is related to a number of explanatory variables $x_2, x_3, \ldots, x_K$ through a linear equation that can be written as:

Eq. 5.3

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_K x_K + e$$

- A single parameter, call it β_k, measures the effect of a change in the variable x_k upon the expected value of y, all other variables held constant

$$\beta_k = \left.\frac{\Delta E(y)}{\Delta x_k}\right|_{\text{other xs held constant}} = \frac{\partial E(y)}{\partial x_k}$$

- The parameter β_1 is the intercept term.
  - We can think of it as being attached to a variable x_1 that is always equal to 1
  - That is, $x_1 = 1$

- The equation for sales revenue can be viewed as a special case of Eq. 5.3 where $K = 3$, $y = SALES$, $x_1 = 1$, $x_2 = PRICE$ and $x_3 = ADVERT$

Eq. 5.4

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

- We make assumptions similar to those we made before:
  - $E(e) = 0$
  - $var(e) = \sigma^2$
    - Errors with this property are said to be **homoskedastic**
  - $cov(e_i, e_j) = 0$
  - $e \sim N(0, \sigma^2)$

- The statistical properties of *y* follow from those of *e*:
  - $E(y) = \beta_1 + \beta_2 x_2 + \beta_3 x_3$
  - $\text{var}(y) = \text{var}(e) = \sigma^2$
  - $\text{cov}(y_i, y_j) = \text{cov}(e_i, e_j) = 0$
  - $y \sim N[(\beta_1 + \beta_2 x_2 + \beta_3 x_3), \sigma^2]$
    - This is equivalent to assuming that $e \sim N(0, \sigma^2)$

- We make two assumptions about the explanatory variables:
  1. The explanatory variables are not random variables
     - We are assuming that the values of the explanatory variables are known to us prior to our observing the values of the dependent variable

- We make two assumptions about the explanatory variables (Continued):
  2. Any one of the explanatory variables is not an exact linear function of the others
     - This assumption is equivalent to assuming that no variable is redundant
     - If this assumption is violated – a condition called **exact collinearity** - the least squares procedure fails

ASSUMPTIONS of the Multiple Regression Model

MR1. $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i, \; i = 1, \ldots, N$

MR2. $E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} \Leftrightarrow E(e_i) = 0$

MR3. $\text{var}(y_i) = \text{var}(e_i) = \sigma^2$

MR4. $\text{cov}(y_i, y_j) = \text{cov}(e_i, e_j) = 0$

MR5. The values of each x_{tk} are not random and are not exact linear functions of the other explanatory variables

MR6. $y_i \sim N\left[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2\right] \Leftrightarrow e_i \sim N(0, \sigma^2)$

5.2
Estimating the Parameters of the Multiple Regression Model

- We will discuss estimation in the context of the model in Eq. 5.4, which we repeat here for convenience, with *i* denoting the *i*th observation

Eq. 5.4

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$$

  - This model is simpler than the full model, yet all the results we present carry over to the general case with only minor modifications

- Mathematically we minimize the sum of squares function $S(\beta_1, \beta_2, \beta_3)$, which is a function of the unknown parameters, given the data:

$$S(\beta_1,\beta_2,\beta_3) = \sum_{i=1}^{N}\left(y_i - E(y_i)\right)^2$$

$$= \sum_{i=1}^{N}\left(y_i - \beta_1 - \beta_2 x_{i2} - \beta_3 x_{i3}\right)^2$$

Eq. 5.5

- Formulas for b_1, b_2, and b_3, obtained by minimizing Eq. 5.5, are estimation procedures, which are called the **least squares estimators** of the unknown parameters
  - In general, since their values are not known until the data are observed and the estimates calculated, the least squares estimators are random variables

- Estimates along with their standard errors and the equation's R^2 are typically reported in equation format as:

Eq. 5.6

$$\widehat{SALES} = 118.91 - 7.908\,PRICE + 1863\,ADVERT \qquad R^2 = 0.448$$
$$(se) \qquad (6.35) \quad (1.096) \qquad (0.683)$$

Table 5.2 Least Squares Estimates for Sales Equation for Big Andy's Burger Barn

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
C	118.9136	6.3516	18.7217	0.0000
PRICE	−7.9079	1.0960	−7.2152	0.0000
ADVERT	1.8626	0.6832	2.7263	0.0080
$R^2 = 0.4483$	$SSE = 1718.943$	$\hat{\sigma} = 4.8861$	$s_y = 6.48854.$	

- Interpretations of the results:
  1. The negative coefficient on *PRICE* suggests that demand is price elastic; we estimate that, with advertising held constant, an increase in price of $1 will lead to a fall in monthly revenue of $7,908
  2. The coefficient on advertising is positive; we estimate that with price held constant, an increase in advertising expenditure of $1,000 will lead to an increase in sales revenue of $1,863

- Interpretations of the results (Continued):
  3. The estimated intercept implies that if both price and advertising expenditure were zero the sales revenue would be $118,914
     - Clearly, this outcome is not possible; a zero price implies zero sales revenue
     - In this model, as in many others, it is important to recognize that the model is an approximation to reality in the region for which we have data
     - Including an intercept improves this approximation even when it is not directly interpretable

- Using the model to predict sales if price is \$5.50 and advertising expenditure is \$1,200:

$$SALES = 118.91 - 7.908PRICE + 1.863ADVERT$$
$$= 118.914 - 7.9079 \times 5.5 + 1.8626 \times 1.2$$
$$= 77.656$$

  - The predicted value of sales revenue for $PRICE = 5.5$ and $ADVERT = 1.2$ is \$77,656.

- A word of caution is in order about interpreting regression results:
  - The negative sign attached to price implies that reducing the price will increase sales revenue.
  - If taken literally, why should we not keep reducing the price to zero?
  - Obviously that would not keep increasing total revenue
  - This makes the following important point:
    - Estimated regression models describe the relationship between the economic variables for values similar to those found in the sample data
    - Extrapolating the results to extreme values is generally not a good idea
    - Predicting the value of the dependent variable for values of the explanatory variables far from the sample values invites disaster

- We need to estimate the error variance, σ^2
  - Recall that:

$$\sigma^2 = \text{var}(e_i) = E\left(e_i^2\right)$$

  - But, the squared errors are unobservable, so we develop an estimator for σ^2 based on the squares of the least squares residuals:

$$\hat{e}_i = y_i - \hat{y}_i = y_i - \left(b_1 + b_2 x_{i2} + b_3 x_{i3}\right)$$

- An estimator for σ^2 that uses the information from $\hat{e}_i^2$ and has good statistical properties is:

Eq. 5.7

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} \hat{e}_i^2}{N - K}$$

where K is the number of β parameters being estimated in the multiple regression model.

- For the hamburger chain example:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{75} \hat{e}_i^2}{N-K} = \frac{1718.943}{75-3} = 23.874$$

- Note that:

$$SSE = \sum_{i=1}^{N} \hat{e}_i^2 = 1718.943$$

  - Also, note that

$$\hat{\sigma} = \sqrt{23.874} = 4.8861$$

  - Both quantities typically appear in the output from your computer software
    - Different software refer to it in different ways.

5.3
Sampling Properties of the Least Squares Estimators

THE GAUSS-MARKOV THEOREM

For the multiple regression model, if assumptions MR1–MR5 hold, then the least squares estimators are the best linear unbiased estimators (*BLUE*) of the parameters.

- If the errors are not normally distributed, then the least squares estimators are approximately normally distributed in large samples
  - What constitutes "large" is tricky
  - It depends on a number of factors specific to each application
  - Frequently, $N - K = 50$ will be large enough

- We can show that:

Eq. 5.8

$$\text{var}(b_2) = \frac{\sigma^2}{(1-r_{23}^2)\sum_{i=1}^{N}(x_{i2}-\bar{x}_2)^2}$$

where

Eq. 5.9

$$r_{23} = \frac{\sum(x_{i2}-\bar{x}_2)(x_{i3}-\bar{x}_3)}{\sqrt{\sum(x_{i2}-\bar{x}_2)^2\sum(x_{i3}-\bar{x}_3)^2}}$$

- We can see that:
  1. Larger error variances σ^2 lead to larger variances of the least squares estimators
  2. Larger sample sizes N imply smaller variances of the least squares estimators
  3. More variation in an explanatory variable around its mean, leads to a smaller variance of the least squares estimator
  4. A larger correlation between x_2 and x_3 leads to a larger variance of b_2. When the correlation is high, it is difficult to disentangle their separate effects.

- We can arrange the variances and covariances in a matrix format:

$$\operatorname{cov}(b_1, b_2, b_3) = \begin{bmatrix} \operatorname{var}(b_1) & \operatorname{cov}(b_1, b_2) & \operatorname{cov}(b_1, b_3) \\ \operatorname{cov}(b_1, b_2) & \operatorname{var}(b_2) & \operatorname{cov}(b_2, b_3) \\ \operatorname{cov}(b_1, b_3) & \operatorname{cov}(b_2, b_3) & \operatorname{var}(b_3) \end{bmatrix}$$

- Using the hamburger data:

Eq. 5.10

$$\widehat{\operatorname{cov}}\left(b_1, b_2, b_3\right) = \begin{bmatrix} 40.343 & -6.795 & -0.7484 \\ -6.795 & 1.201 & -0.0197 \\ -0.7484 & -0.0197 & 0.4668 \end{bmatrix}$$

- We are particularly interested in the standard errors:

$$se(b_1) = \sqrt{\widehat{\text{var}}(b_1)} = \sqrt{40.343} = 6.3516$$

$$se(b_2) = \sqrt{\widehat{\text{var}}(b_2)} = \sqrt{1.201} = 1.0960$$

$$se(b_3) = \sqrt{\widehat{\text{var}}(b_3)} = \sqrt{0.4668} = 0.6832$$

Table 5.3 Covariance Matrix for Coefficient Estimates

	C	*PRICE*	*ADVERT*
C	40.3433	−6.7951	−0.7484
PRICE	−6.7951	1.2012	−0.0197
ADVERT	−0.7484	−0.0197	0.4668

- Consider the general for of a multiple regression model:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_K x_{iK} + e_i$$

  - If we add assumption MR6, that the random errors e_i have normal probability distributions, then the dependent variable y_i is normally distributed:

$$y_i \sim N\left[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2\right] \Leftrightarrow e_i \sim N(0, \sigma^2)$$

- Since the least squares estimators are linear functions of dependent variables, it follows that the least squares estimators are also normally distributed:

$$b_k \sim N\left[\beta_k, \text{var}\left(b_k\right)\right]$$

- We can now form the standard normal variable Z:

Eq. 5.11

$$z = \frac{b_k - \beta_k}{\sqrt{\text{var}(b_k)}} \sim N(0,1), \text{ for } k = 1, 2, \ldots, K$$

- Replacing the variance of b_k with its estimate:

Eq. 5.12

$$t = \frac{b_k - \beta_k}{\sqrt{\widehat{\text{var}}(b_k)}} = \frac{b_k - \beta_k}{\text{se}(b_k)} \sim t_{(N-K)}$$

  - Notice that the number of degrees of freedom for *t*-statistics is $N - K$

- We can form a linear combination of the coefficients as:

$$\lambda = c_1\beta_1 + c_2\beta_2 + \ldots + c_K\beta_K = \sum_{k=1}^{K} c_k\beta_k$$

- And then we have

$$t = \frac{\hat{\lambda} - \lambda}{se(\hat{\lambda})} = \frac{\sum c_k b_k - \sum c_k\beta_k}{\text{se}(\sum c_k b_k)} \sim t_{(N-K)}$$

Eq. 5.13

- If $K = 3$, the we have:

$$se(c_1 b_1 + c_2 b_2 + c_3 b_3) = \sqrt{\widehat{var}(c_1 b_1 + c_2 b_2 + c_3 b_3)}$$

where

$$\widehat{var}(c_1 b_1 + c_2 b_2 + c_3 b_3) = c_1^2 \widehat{var}(b_1) + c_2^2 \widehat{var}(b_2) + c_3^2 \widehat{var}(b_3)$$

Eq. 5.14

$$+2c_1 c_2 \widehat{cov}(b_1, b_2) + 2c_1 c_3 \widehat{cov}(b_1, b_3) + 2c_2 c_3 \widehat{cov}(b_2, b_3)$$

- What happens if the errors are not normally distributed?
  - Then the least squares estimator will not be normally distributed and Eq. 5.11, Eq. 5.12, and Eq. 5.13 will not hold exactly
  - They will, however, be approximately true in large samples
  - Thus, having errors that are not normally distributed does not stop us from using Eq. 5.12 and Eq. 5.13, but it does mean we have to be cautious if the sample size is not large
  - A test for normally distributed errors was given in Chapter 4.3.5

5.4
Interval Estimation

- For the hamburger example, we need:

Eq. 5.15

$$P(-t_c < t_{(72)} < t_c) = .95$$

- Using $t_c = 1.993$, we can rewrite (5.15) as:

$$P\left(-1.993 \leq \frac{b_2 - \beta_2}{\text{se}(b_2)} \leq 1.993\right) = .95$$

- Rearranging, we get:

$$P\left[b_2 - 1.993 \times \text{se}(b_2) \leq \beta_2 \leq b_2 + 1.993 \times \text{se}(b_2)\right] = .95$$

  - Or, just writing the end-points for a 95% interval:

$$\left[b_2 - 1.993 \times \text{se}(b_2),\ b_2 + 1.993 \times \text{se}(b_2)\right]$$

Eq. 5.16

- Using our data, we have $b_2 = -7.908$ and $se(b_2) = 1.096$, so that:

$$(-7.9079 - 1.993 \times 1.096, -7.9079 + 1.993 \times 1.096) = (-10.093, -5.723)$$

  - This interval estimate suggests that decreasing price by $1 will lead to an increase in revenue somewhere between $5,723 and $10,093.
    - In terms of a price change whose magnitude is more realistic, a 10-cent price reduction will lead to a revenue increase between $572 and $1,009

- Similarly for advertising, we get:

$$(1.8626 - 1.9935 \times 0.6832, 1.8626 + 1.9935 \times 0.6832) = (0.501, 3.225)$$

  - We estimate that an increase in advertising expenditure of $1,000 leads to an increase in sales revenue of between $501 and $3,225
  - This interval is a relatively wide one; it implies that extra advertising expenditure could be unprofitable (the revenue increase is less than $1,000) or could lead to a revenue increase more than three times the cost of the advertising

- We write the general expression for a 100(1-α)% confidence interval as:

$$\left(b_k - t_{(\alpha/2,\, N-K)} \times \text{se}(b_k),\ b_k + t_{(1-\alpha/2,\, N-K)} \times \text{se}(b_k)\right)$$

- Suppose Big Andy wants to increase advertising expenditure by \$800 and drop the price by 40 cents.
  - Then the change in expected sales is:

$$\begin{aligned}\lambda &= E(SALES_1) - E(SALES_0) \\ &= \left[\beta_1 + \beta_2(PRICE_0 - 0.4) + \beta_3(ADVERT_0 + 0.8)\right] \\ &\quad -\left[\beta_1 + \beta_2 PRICE_0 + \beta_3 ADVERT_0\right] \\ &= -0.4\beta_2 + 0.8\beta_3\end{aligned}$$

- A point estimate would be:

$$\hat{\lambda} = -0.4b_2 + 0.8b_3 = -0.4 \times (-7.9079) + 0.8 \times 1.8626$$
$$= 4.6532$$

- A 90% interval would be:

$$\left(\hat{\lambda} - t_c \times se\left(\hat{\lambda}\right), \hat{\lambda} - t_c \times se\left(\hat{\lambda}\right)\right)$$
$$= \left(\left(-0.4b_2 + 0.8b_3\right) - t_c \times se\left(-0.4b_2 + 0.8b_3\right),\right.$$
$$\left.\left(-0.4b_2 + 0.8b_3\right) + t_c \times se\left(-0.4b_2 + 0.8b_3\right)\right)$$

- The standard error is:

$$\begin{aligned}
\text{se}\left(-0.4b_2+0.8b_3\right) &= \sqrt{\text{var}\left(-0.4b_2+0.8b_3\right)} \\
&= \sqrt{\left(-0.4\right)^2\widehat{\text{var}\left(b_2\right)}+\left(0.8\right)^2\widehat{\text{var}\left(b_3\right)}-2\times-0.4\times0.8\times\widehat{\text{cov}\left(b_2,b_3\right)}} \\
&= \sqrt{0.16\times1.2012+0.64\times0.4668-0.64\times\left(-0.0197\right)} \\
&= 0.7096
\end{aligned}$$

- The 90% interval is then:

$$(4.6532 - 1.666 \times 0.7096, 4.6532 + 1.666 \times 0.7096) = (3.471, 5.835)$$

  - We estimate, with 90% confidence, that the expected increase in sales will lie between \$3,471 and \$5,835

5.5
Hypothesis Testing

COMPONENTS OF HYPOTHESIS TESTS

1. A null hypothesis H_0
2. An alternative hypothesis H_1
3. A test statistic
4. A rejection region
5. A conclusion

- We need to ask whether the data provide any evidence to suggest that y is related to each of the explanatory variables
  - If a given explanatory variable, say x_k, has no bearing on y, then $\beta_k = 0$
  - Testing this null hypothesis is sometimes called a **test of significance** for the explanatory variable x_k

- Null hypothesis:

$$H_0 : \beta_k = 0$$

- Alternative hypothesis:

$$H_1 : \beta_k \neq 0$$

- Test statistic:

$$t = \frac{b_k}{\text{se}(b_k)} \sim t_{(N-K)}$$

- critical values for a test with level of significance α:

$$t_c = t_{(1-\alpha/2,\, N-K)} \text{ and } -t_c = t_{(\alpha/2,\, N-K)}$$

- For our hamburger example, we can conduct a test that sales revenue is related to price:

1. The null and alternative hypotheses are:

$$H_0 : \beta_2 = 0 \text{ and } H_1 : \beta_2 \neq 0$$

2. The test statistic, if the null hypothesis is true, is:

$$t = b_2 / \text{se}(b_2) \sim t_{(N-K)}$$

3. Using a 5% significance level (α=.05), and 72 degrees of freedom, the critical values that lead to a probability of 0.025 in each tail of the distribution are:

$$t_{(.975,72)} = 1.993 \text{ and } t_{(.025,72)} = -1.993$$

- For our hamburger example (Continued) :

4. The computed value of the *t*-statistic is:

$$t = \frac{-7.908}{1.096} = -7.215$$

and the *p*-value from software is:

$$P\left(t_{(72)} > 7.215\right) + P\left(t_{(72)} < -7.215\right) = 2 \times (2.2 \times 10^{-10}) = 0.000$$

5. Since -7:215 < -1.993, we reject H_0: $\beta_2 = 0$ and conclude that there is evidence from the data to suggest sales revenue depends on price
   - Using the *p*-value to perform the test, we reject H_0 because $0.000 < 0.05$.

- Similarly, we can conduct a test that sales revenue is related to advertising expenditure:

1. The null and alternative hypotheses are:

$$H_0 : \beta_3 = 0 \text{ and } H_1 : \beta_3 \neq 0$$

2. The test statistic, if the null hypothesis is true, is:

$$t = b_3 / \text{se}(b_3) \sim t_{(N-K)}$$

3. Using a 5% significance level (α=.05), and 72 degrees of freedom, the critical values that lead to a probability of 0.025 in each tail of the distribution are:

$$t_{(.975,72)} = 1.993 \text{ and } t_{(.025,72)} = -1.993$$

- For our hamburger example (Continued):

  4. The computed value of the *t*-statistic is:

$$t = \frac{1.8626}{0.6832} = 2.726$$

  and the *p*-value from software is:

$$P\left(t_{(72)} > 2.726\right) + P\left(t_{(72)} < -2.726\right) = 2 \times 0.004 = 0.008$$

  5. Since $2.726 > 1.993$, we reject H_0: $\beta_3 = 0$: the data support the conjecture that revenue is related to advertising expenditure
     - Using the *p*-value to perform the test, we reject H_0 because $0.008 < 0.05$

- However, do not confuse statistical significance with economic importance and precision.

- We now are in a position to state the following questions as testable hypotheses and ask whether the hypotheses are compatible with the data
  1. Is demand price-elastic or price-inelastic?
  2. Would additional sales revenue from additional advertising expenditure cover the costs of the advertising?

- For the demand elasticity, we wish to know if:
  - $\beta_2 \geq 0$: an increase in price leads to an increase in sales revenue (demand is price-inelastic or has an elasticity of unity), or
  - $\beta_2 < 0$: an increase in price leads to a decrease in sales revenue (demand is price-elastic)

- As before:

1. The null and alternative hypotheses are:

$H_0 : \beta_2 \geq 0$ (demand is unit-elastic or inelastic)

$H_1 : \beta_2 < 0$ (demand is elastic)

2. The test statistic, if the null hypothesis is true, is:

$$t = b_2 / \text{se}(b_2) \sim t_{(N-K)}$$

3. At a 5% significance level, we reject H_0 if $t \leq -1.666$ or if the p-value ≤ 0.05

- Hypothesis test (Continued) :

4. The test statistic is:

$$t = \frac{b_2}{se(b_2)} = \frac{-7.908}{1.096} = -7.215$$

and the *p*-value is:

$$P\left(t_{(72)} < -7.215\right) = 0.000$$

5. Since -7.215 < -1.666, we reject H_0: $\beta_2 \geq 0$ and conclude that H_0: $\beta_2 < 0$ (demand is elastic)

- The other hypothesis of interest is whether an increase in advertising expenditure will bring an increase in sales revenue that is sufficient to cover the increased cost of advertising
  - Such an increase will be achieved if $\beta_3 > 1$

- As before:

1. The null and alternative hypotheses are:

$$H_0 : \beta_3 \leq 1$$

$$H_1 : \beta_3 > 1$$

2. The test statistic, if the null hypothesis is true, is:

$$t = \frac{b_3 - 1}{\text{se}(b_3)} \sim t_{(N-K)}$$

3. At a 5% significance level, we reject H_0 if $t \geq 1.666$ or if the p-value ≤ 0.05

- Hypothesis test (Continued) :

4. The test statistic is:

$$t = \frac{b_3 - \beta_3}{se(b_2)} = \frac{1.8626 - 1}{0.6832} = 1.263$$

and the *p*-value is:

$$P\left(t_{(72)} > 1.263\right) = 0.105$$

5. Since $1.263 < 1.666$, we do not reject H_0

- The marketing adviser claims that dropping the price by 20 cents will be more effective for increasing sales revenue than increasing advertising expenditure by $500
  - In other words, she claims that $-0.2\beta_2 > 0.5\beta_3$, or $-0.2\beta_2 - 0.5\beta_3 > 0$
  - We want to test a hypothesis about the linear combination $-0.2\beta_2 - 0.5\beta_3$

- As before:

1. The null and alternative hypotheses are:

$H_0 : -0.2\beta_2 - 0.5\beta_3 \leq 0$ (marketer's claim is not correct)

$H_1 : -0.2\beta_2 - 0.5\beta_3 > 0$ (marketer's claim is correct)

2. The test statistic, if the null hypothesis is true, is: $$t = \frac{-0.2b_2 - 0.5b_3}{\text{se}(-0.2b_2 - 0.5b_3)} \sim t_{(72)}$$

3. At a 5% significance level, we reject H_0 if $t \geq 1.666$ or if the p-value ≤ 0.05

- We need the standard error:

$$\text{se}(-0.2b_2 - 0.5b_3) = \sqrt{\widehat{\text{var}}\left(\text{se}(-0.2b_2 - 0.5b_3)\right)}$$

$$= \sqrt{(-0.2)^2\,\widehat{\text{var}}(b_2) + (-0.5)^2\,\widehat{\text{var}}(b_3) + 2\times(-0.2)\times(-0.5)\,\widehat{\text{cov}}(b_2, b_3)}$$

$$= \sqrt{0.04\times 1.2012 + 0.25\times 0.4668 + 0.2\times(-0.0197)}$$

$$= 0.4010$$

- Hypothesis test (Continued) :

4. The test statistic is:

$$t = \frac{-0.2b_2 - 0.5b_3}{\text{se}\left(-0.2b_2 - 0.5b_3\right)} = \frac{1.58158 - 0.9319}{0.4010} = 1.622$$

and the p-value is:

$$P\left(t_{(72)} > 1.622\right) = 0.055$$

5. Since $1.622 < 1.666$, we do not reject H_0

5.6
Interaction Variables

- In the multiple regression model the marginal effect of each variable is constant and independent on other variables.
- What if we wanted the marginal effect of one variable to depend on the level of the other variable?
- Suppose that we wish to study the effect of income and age on an individual's expenditure on pizza
  - An initial model would be:

Eq. 5.17

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_3 INCOME + e$$

- Implications of this model are:
  1. $\partial E(PIZZA)/\partial AGE = \beta_2$: For a given level of income, the expected expenditure on pizza changes by the amount β_2 with an additional year of age
  2. $\partial E(PIZZA)/\partial INCOME = \beta_3$: For individuals of a given age, an increase in income of \$1,000 increases expected expenditures on pizza by β_3

Table 5.4 Pizza Expenditure Data

PIZZA	*INCOME*	*AGE*
109	19.5	25
0	39.0	45
0	15.6	20
108	26.0	28
220	19.5	25

- The estimated model is:

$$PIZZA = 342.88 - 7.576AGE + 1.832INCOME$$

$$(t) \qquad (-3.27)^{***} \qquad (3.95)^{***}$$

  - The signs of the estimated parameters are as we anticipated
    - Both *AGE* and *INCOME* have significant coefficients, based on their *t*-statistics

- It is not reasonable to expect that, regardless of the age of the individual, an increase in income by $1,000 should lead to an increase in pizza expenditure by $1.83.
  - It would seem more reasonable to assume that as a person grows older, his or her marginal propensity to spend on pizza declines
    - That is, as a person ages, less of each extra dollar is expected to be spent on pizza
  - This is a case in which the effect of income depends on the age of the individual.
    - That is, the effect of one variable is modified by another
  - One way of accounting for such interactions is to include an **interaction variable** that is the product of the two variables involved

- We will add the interaction variable (*AGE* x *INCOME*) to the regression model
  - The new model is:

Eq. 5.18

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_3 INCOME + \beta_4 (AGE \times INCOME) + e$$

- Implications of this revised model are:
  1. $\partial E(PIZZA)/\partial AGE = \beta_2 + \beta_4 INCOME$
     As age increases the consumption of pizza depends on income. If $\beta_4 < 0$, then, the larger the income the smaller this change of consumption would be.
  2. $\partial E(PIZZA)/\partial INCOME = \beta_3 + \beta_4 AGE$
     As income increases the consumption of pizza depends on age. If $\beta_4 < 0$, then, as age increases this change of consumption will decrease.

- The estimated model is:

$$PIZZA = 161.47 - 2.977\,AGE + 6.980\,INCOME - 0.1232\left(AGE \times INCOME\right)$$

$$\text{(t)} \qquad (-0.89) \qquad (2.47)^{**} \qquad (-1.85)^{*}$$

- The estimated marginal effect of age upon pizza expenditure for two individuals—one with $25,000 income and one with $90,000 income is:

$$\widehat{\frac{\partial E(PIZZA)}{\partial AGE}} = b_2 + b_4 INCOME$$

$$= -2.977 - 0.1232 INCOME$$

$$= \begin{cases} -6.06 & \text{for } INCOME = 25 \\ -14.07 & \text{for } INCOME = 90 \end{cases}$$

  - We expect that an individual with $25,000 income will reduce pizza expenditures by $6.06 per year, whereas the individual with $90,000 income will reduce pizza expenditures by $14.07 per year

- Consider a wage equation where ln(*WAGE*) depends on years of education (*EDUC*) and years of experience (*EXPER*):

Eq. 5.19

$$\ln(WAGE) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + e$$

  - If we believe the effect of an extra year of experience on wages will depend on the level of education, then we can add an interaction variable

Eq. 5.20

$$\ln(WAGE) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 (EDUC \times EXPER) + e$$

- The effect of another year of experience, holding education constant, is roughly:

$$\left.\frac{\Delta \ln(WAGE)}{\Delta EXPER}\right|_{EDUC \text{ fixed}} = \beta_3 + \beta_4 EDUC$$

  - The approximate percentage change in wage given a one-year increase in experience is $100(\beta_3+\beta_4 EDUC)\%$

- An estimated model is:

$$\widehat{\ln(WAGE)} = 1.392 + 0.09494EDUC + 0.00633EXPER$$
$$-0.0000364(EDUC \times EXPER)$$

- The estimate of β_4 is negative. Thus:
  - An extra year of education has a smaller effect on the wage the greater the number of years of experience is.
  - An extra year of experience has a smaller effect on the wage the greater the number of years of education is.
- For a person with 8 years of education an additional year of experience leads to increase of wage of

$$100(0.00633 - 0.0000364 \times 8)\% = 0.6\%$$

- For a person with 16 years of education an additional year of experience leads to increase of wage of

$$100(0.00633 - 0.0000364 \times 16)\% = 0.57\%$$

5.7
Measuring Goodness-of-fit

- In the multiple regression model the R^2 is relevant and the same formulas are valid, but now we talk of the proportion of variation in the dependent variable explained by all the explanatory variables included in the linear model

- The coefficient of determination is:

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

$$= 1 - \frac{SSE}{SST}$$

$$= 1 - \frac{\sum_{i=1}^{N}\hat{e}_i^{\,2}}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

Eq. 5.21

- The predicted value of y is:

$$\hat{y}_i = \mathrm{b}_1 + \mathrm{b}_2 x_{i2} + b_3 x_{i3} + \cdots + b_K x_{iK}$$

  - Recall that:

$$s_y = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} = \sqrt{\frac{SST}{N-1}}$$

  - Then:

$$SST = \left(N-1\right)s_y^2$$

- For the hamburger example:.

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \hat{e}_i^{\,2}}{\sum_{i=1}^{N} (y_i - \bar{y})^2} 1 - \frac{1718.943}{3115.482} = 0.448$$

- Interpretation
  - 44.8% of the variation in sales revenue is explained by the variation in price and by the variation in the level of advertising expenditure
  - In our sample, 55.2% of the variation in revenue is left unexplained and is due to variation in the error term or to variation in other variables that implicitly form part of the error term

Key Words

Keywords

- BLU estimator
- covariance matrix of least squares estimator
- critical value
- error variance estimate
- error variance estimator
- goodness-of-fit
- interaction variable
- interval estimate
- least squares estimates
- least squares estimation
- least squares estimators
- linear combinations
- marginal effects
- multiple regression model
- Nonlinear functions
- one-tailed test
- *p*-value
- regression coefficients
- standard errors
- sum of squared errors
- sum of squares of regression
- testing significance
- total sum of squares
- two-tailed test