

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

The Coefficient of Determination

Lecture 49 Section 13.9

Robb T. Koether

Hampden-Sydney College

Tue, Apr 22, 2008

Outline

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- 1 Introduction
- 2 The Regression Identity
- 3 Sums of Squares on the TI-83
- 4 Explaining Variation
- 5 The Coefficient of Determination

Introduction

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- One goal of regression is to produce a model that will predict values of the response variable, given values of the explanatory variable.
- We have already done that.
- Another goal is to “explain” the variability in the response variable as a consequence of the variability in the explanatory variable.
- That is where we now turn our attention.

Residual Sum of Squares

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Recall that the line of “best” fit was that line with the smallest sum of squared residuals.
- This is also called the **residual sum of squares** or **sum of squared errors**:

$$\text{SSE} = \sum (y - \hat{y})^2.$$

Other Sums of Squares

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- There are two other sums of squares associated with y .
- The **regression sum of squares**:

$$\text{SSR} = \sum (\hat{y} - \bar{y})^2.$$

- The **total sum of squares**:

$$\text{SST} = \text{SSY} = \sum (y - \bar{y})^2.$$

Example - SST, SSR, and SSE

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Plot the following data.

x	y
1	8
3	12
4	9
5	14
8	16
9	20
11	17
15	24

Example - SST, SSR, and SSE

The
Coefficient of
Determination

Robb T.
Koether

Introduction

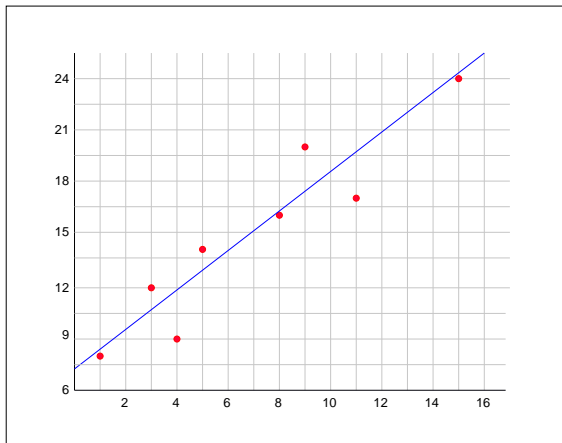
The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

The regression line



Example - SST, SSR, and SSE

The Coefficient of Determination

Robb T. Koether

Introduction

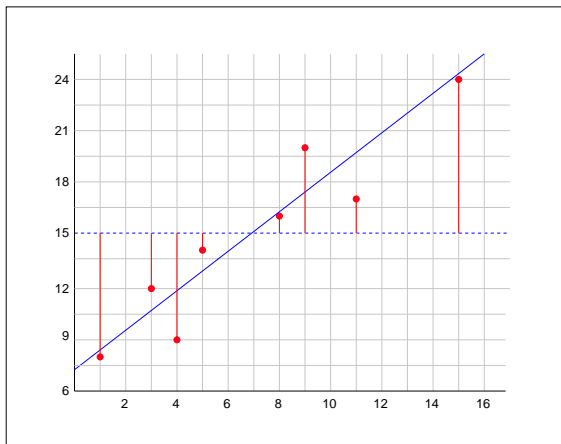
The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

The deviations of y from \bar{y}



Example - SST, SSR, and SSE

The
Coefficient of
Determination

Robb T.
Koether

Introduction

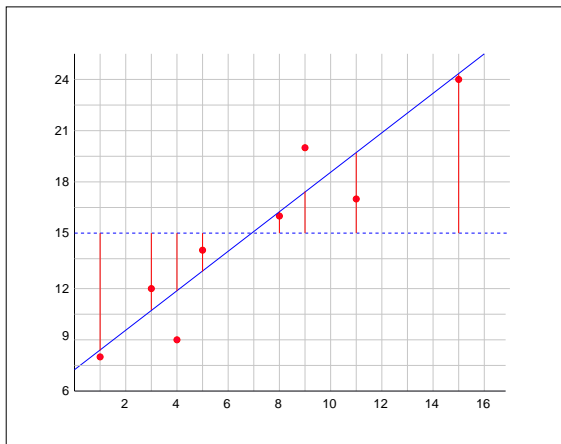
The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

The deviations of \hat{y} from \bar{y}



Example - SST, SSR, and SSE

The
Coefficient of
Determination

Robb T.
Koether

Introduction

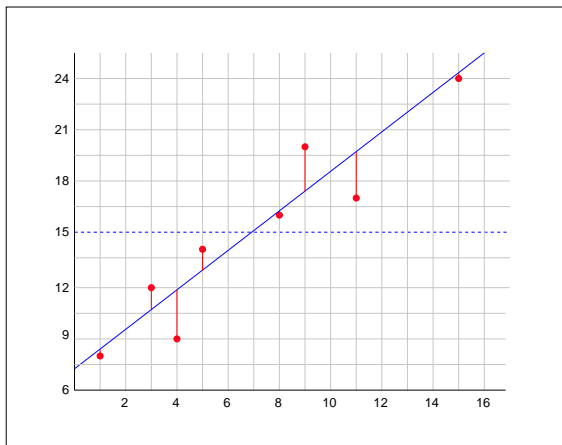
The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

The deviations of y from \hat{y}



The Regression Identity

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- It turns out that

$$SST = SSR + SSE.$$

- This is the **regression identity**.
- It also turns out that

$$r^2 = \frac{SSR}{SST}.$$

- Consequently,

$$r^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST} = 1 - \frac{SSE}{SST}.$$

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SST.

x	y	$y - \bar{y}$	$(y - \bar{y})^2$
1	8		
3	12		
4	9		
5	14		
8	16		
9	20		
11	17		
15	24		

Example

The Coefficient of Determination

Robb T. Koether

Introduction

The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

- Compute SST.

x	y	$y - \bar{y}$	$(y - \bar{y})^2$
1	8	-7	
3	12	-3	
4	9	-6	
5	14	-1	
8	16	1	
9	20	5	
11	17	2	
15	24	9	

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SST.

x	y	$y - \bar{y}$	$(y - \bar{y})^2$
1	8	-7	49
3	12	-3	9
4	9	-6	36
5	14	-1	1
8	16	1	1
9	20	5	25
11	17	2	4
15	24	9	81

Example

The Coefficient of Determination

Robb T. Koether

Introduction

The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

- Compute SST.

x	y	$y - \bar{y}$	$(y - \bar{y})^2$
1	8	-7	49
3	12	-3	9
4	9	-6	36
5	14	-1	1
8	16	1	1
9	20	5	25
11	17	2	4
15	24	9	81
			206

Example

The Coefficient of Determination

Robb T. Koether

Introduction

The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

- Compute SSR.

x	y	\hat{y}	$\hat{y} - \bar{y}$	$(\hat{y} - \bar{y})^2$
1	8	8.4		
3	12	10.6		
4	9	11.7		
5	14	12.8		
8	16	16.1		
9	20	17.2		
11	17	19.4		
15	24	23.8		

Example

The Coefficient of Determination

Robb T. Koether

Introduction

The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

- Compute SSR.

x	y	\hat{y}	$\hat{y} - \bar{y}$	$(\hat{y} - \bar{y})^2$
1	8	8.4	-6.6	
3	12	10.6	-4.4	
4	9	11.7	-3.3	
5	14	12.8	-2.2	
8	16	16.1	1.1	
9	20	17.2	2.2	
11	17	19.4	4.4	
15	24	23.8	8.8	

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SSR.

x	y	\hat{y}	$\hat{y} - \bar{y}$	$(\hat{y} - \bar{y})^2$
1	8	8.4	-6.6	43.56
3	12	10.6	-4.4	19.36
4	9	11.7	-3.3	10.89
5	14	12.8	-2.2	4.84
8	16	16.1	1.1	1.21
9	20	17.2	2.2	4.84
11	17	19.4	4.4	19.36
15	24	23.8	8.8	77.44

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SSR.

x	y	\hat{y}	$\hat{y} - \bar{y}$	$(\hat{y} - \bar{y})^2$
1	8	8.4	-6.6	43.56
3	12	10.6	-4.4	19.36
4	9	11.7	-3.3	10.89
5	14	12.8	-2.2	4.84
8	16	16.1	1.1	1.21
9	20	17.2	2.2	4.84
11	17	19.4	4.4	19.36
15	24	23.8	8.8	77.44
				181.50

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SSE.

x	y	\hat{y}	$y - \hat{y}$	$(y - \hat{y})^2$
1	8	8.4		
3	12	10.6		
4	9	11.7		
5	14	12.8		
8	16	16.1		
9	20	17.2		
11	17	19.4		
15	24	23.8		

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SSE.

x	y	\hat{y}	$y - \hat{y}$	$(y - \hat{y})^2$
1	8	8.4	-0.4	
3	12	10.6	1.4	
4	9	11.7	-1.3	
5	14	12.8	1.2	
8	16	16.1	-0.1	
9	20	17.2	2.8	
11	17	19.4	-2.4	
15	24	23.8	0.2	

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SSE.

x	y	\hat{y}	$y - \hat{y}$	$(y - \hat{y})^2$
1	8	8.4	-0.4	0.16
3	12	10.6	1.4	1.96
4	9	11.7	-1.3	1.69
5	14	12.8	1.2	1.44
8	16	16.1	-0.1	0.01
9	20	17.2	2.8	7.84
11	17	19.4	-2.4	5.76
15	24	23.8	0.2	0.04

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Compute SSE.

x	y	\hat{y}	$y - \hat{y}$	$(y - \hat{y})^2$
1	8	8.4	-0.4	0.16
3	12	10.6	1.4	1.96
4	9	11.7	-2.7	7.29
5	14	12.8	1.2	1.44
8	16	16.1	-0.1	0.01
9	20	17.2	2.8	7.84
11	17	19.4	-2.4	5.76
15	24	23.8	0.2	0.04
				24.50

Example

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- We have now found that

$$SST = 206.$$

$$SSR = 181.50.$$

$$SSE = 24.50.$$

- We can verify that $SSR + SSE = SST$:

$$181.50 + 24.50 = 206.$$

TI-83 - Finding SST, SSR, and SSE

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Put the x values into L_1 and the y values into L_2 .
- Use $\text{LinReg}(a+bx)$ L_1, L_2, Y_1 .
- Enter $Y_1(L_1) \rightarrow L_3$.
- To get SST,
 - Use 1-Var Stats L_2 to get \bar{y} .
 - Evaluate $\text{sum}(L_2 - \bar{y})^2$.
- To get SSR, evaluate $\text{sum}(L_3 - \bar{y})^2$.
- To get SSE, evaluate $\text{sum}((L_2 - L_3)^2)$.

TI-83 - Finding SST, SSR, and SSE

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Use the TI-83 to find SST, SSR, and SSE in this example.

x	y
1	8
3	12
4	9
5	14
8	16
9	20
11	17
15	24

Explaining Variation

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- One goal of regression is to “explain” the variation in y .
- For example, if y were weight, how would we explain the variation in weight?
- That is, why do some people weigh more than others?
- A *partial* answer is that some people weigh more because they are taller.
- That is, an explanatory variable is height x .
- What are some other partial answers?

Explaining Variation

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- *How much* of the variation in weight is explained by variation in height?
- The total variation in weight is SST.
- The linear model (the regression line) explains some of the variation.
- The model predicts the variation SSR.
- The remainder is SSE, the variation not predicted by the model.

Explaining Variation

The
Coefficient of
Determination

Robb T.
Koether

Introduction

The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

- Statisticians consider the predicted variation SSR to be the amount of variation in y that is explained by the model.
- The residual variation SSE is the remaining variation in y that is not explained by the model.
- It all checks out because $SST = SSR + SSE$.

Variation Explained by the Model

The Coefficient of Determination

Robb T. Koether

Introduction

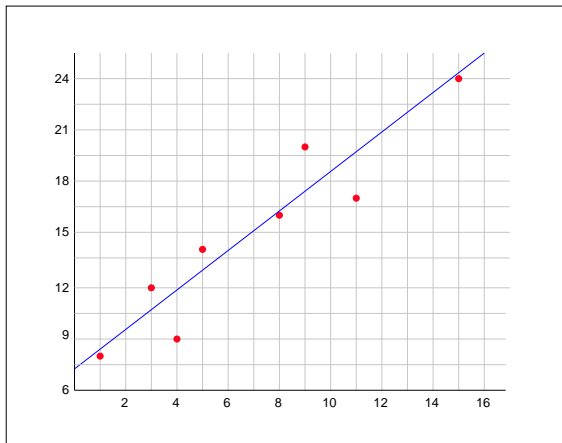
The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

The regression line



Variation Explained by the Model

The Coefficient of Determination

Robb T. Koether

Introduction

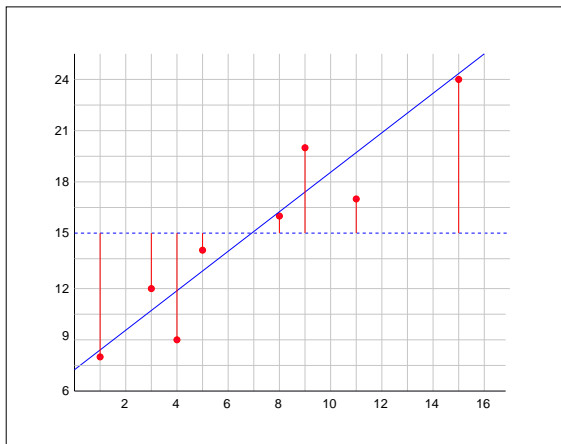
The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

The total variation in y (SST)



Variation Explained by the Model

The
Coefficient of
Determination

Robb T.
Koether

Introduction

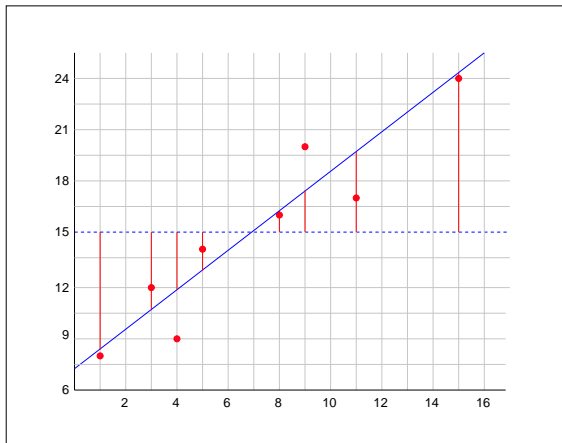
The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

The variation in y that is explained by the model (SSR)



Variation Explained by the Model

The
Coefficient of
Determination

Robb T.
Koether

Introduction

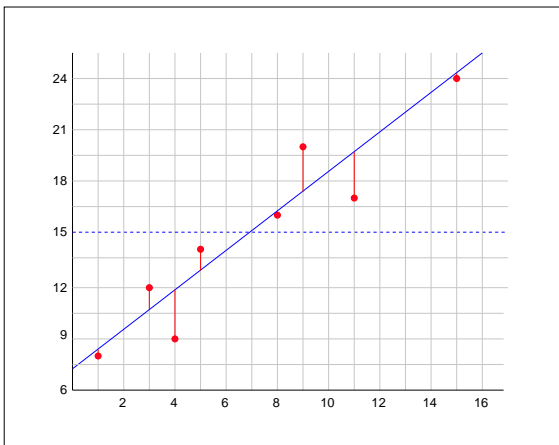
The
Regression
Identity

Sums of
Squares on
the TI-83

Explaining
Variation

The
Coefficient of
Determination

The variation in y that is unexplained by the model (SSE)



Explaining Variation

The Coefficient of Determination

Robb T. Koether

Introduction

The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

- Recall that

$$r^2 = \frac{SSR}{SST}$$

and

$$1 - r^2 = \frac{SSE}{SST}.$$

- Therefore, r^2 is the proportion of variation in y that is explained by the model. It is called the **coefficient of determination**.
- $1 - r^2$ is the proportion that is not explained by the model.

TI-83 - Calculating r^2

The Coefficient of Determination

Robb T.
Koether

Introduction

The Regression Identity

Sums of Squares on the TI-83

Explaining Variation

The Coefficient of Determination

- To calculate r^2 on the TI-83, follow the procedure that produces the regression line and r .
- In the same window, the TI-83 reports the value of r^2 .