

Paired Samples

Lecture 37

Sections 11.1, 11.2, 11.3

Robb T. Koether

Hampden-Sydney College

Fri, Mar 30, 2012

Outline

- 1 Introduction
- 2 Independent vs. Dependent Samples
- 3 Dependent Samples (Paired Data)
- 4 Assignment

Outline

- 1 Introduction
- 2 Independent vs. Dependent Samples
- 3 Dependent Samples (Paired Data)
- 4 Assignment

Introduction

- In this chapter we will consider problems that compare two populations.

Introduction

- In this chapter we will consider problems that compare two populations.
- For example, we could compare

Introduction

- In this chapter we will consider problems that compare two populations.
- For example, we could compare
 - The proportion of men who vote Republican vs. the proportion of women who vote Republican. (p_1 vs. p_2 .)

Introduction

- In this chapter we will consider problems that compare two populations.
- For example, we could compare
 - The proportion of men who vote Republican vs. the proportion of women who vote Republican. (p_1 vs. p_2 .)
 - The average age at which a male first marries vs. the average age at which a female first marries. (μ_1 vs. μ_2 .)

Introduction

- In this chapter we will consider problems that compare two populations.
- For example, we could compare
 - The proportion of men who vote Republican vs. the proportion of women who vote Republican. (p_1 vs. p_2 .)
 - The average age at which a male first marries vs. the average age at which a female first marries. (μ_1 vs. μ_2 .)
- The bad news: Things will get a bit more complicated.

Introduction

- In this chapter we will consider problems that compare two populations.
- For example, we could compare
 - The proportion of men who vote Republican vs. the proportion of women who vote Republican. (p_1 vs. p_2 .)
 - The average age at which a male first marries vs. the average age at which a female first marries. (μ_1 vs. μ_2 .)
- The bad news: Things will get a bit more complicated.
- The good news: You will not need to memorize any more formulas.

Outline

- 1 Introduction
- 2 Independent vs. Dependent Samples**
- 3 Dependent Samples (Paired Data)
- 4 Assignment

Dependent Samples

Definition (Bivariate data, paired samples)

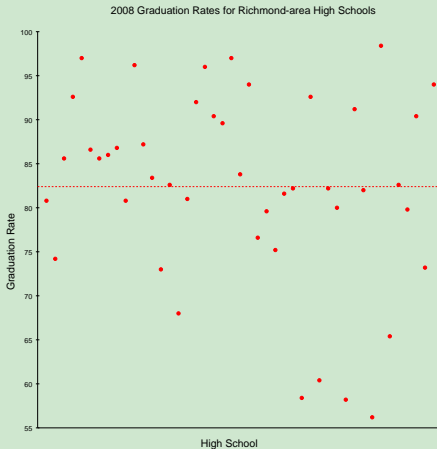
Bivariate data are data in which each datum is a pair of observations. These are also called **paired data**. Typically the two values are called x_1 and x_2 . The sample of x_1 values and the sample of x_2 values are called **paired samples** or **dependent samples**.

Dependent Samples

- Paired data are often “before” and “after” observations.
- By comparing the mean before treatment to the mean after treatment, we can determine whether the treatment had an effect.
- To make direct comparisons of the two samples, they must be measuring the same sort of thing.
- Clearly, paired samples must be of the same size.

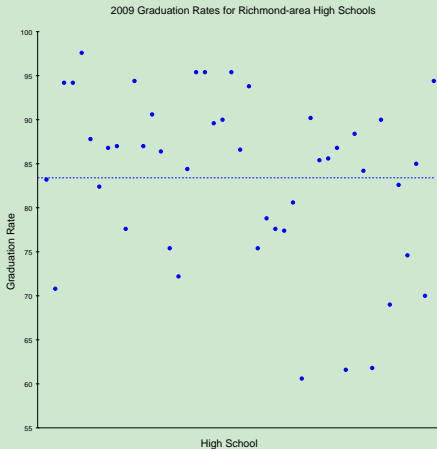
Dependent Samples

Example (High-School Graduation Rates)



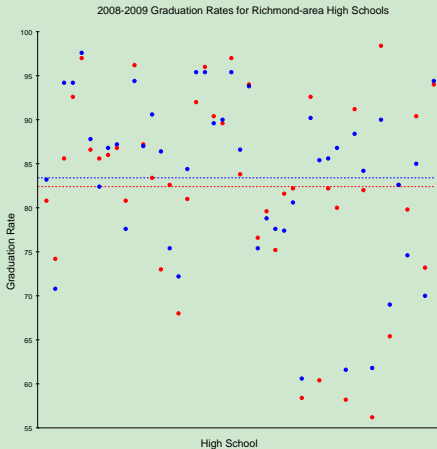
Dependent Samples

Example (High-School Graduation Rates)



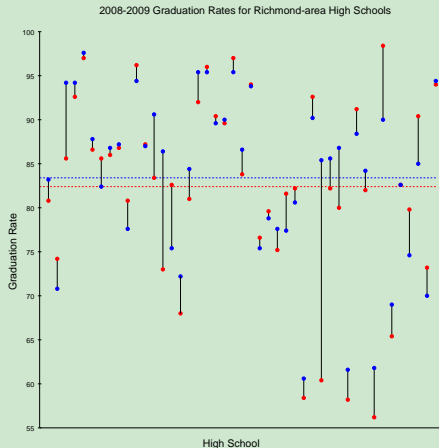
Dependent Samples

Example (High-School Graduation Rates)



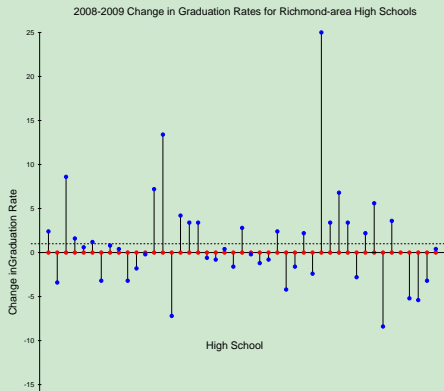
Dependent Samples

Example (High-School Graduation Rates)



Dependent Samples

Example (High-School Graduation Rates)



Dependent Samples

Example (High-School Graduation Rates)

- Was there an overall improvement in the graduation rate?
- That is, is the average difference greater than 0?
- See the article [Va. graduation, dropout rates improve over last year.](#)

Independent Samples

- On the other hand, with **independent samples**, we simply take one sample from one population and another sample from another population.
- There is no logical way to pair the data.
- Furthermore, the independent samples could be of different sizes.
- In this chapter, we will first study paired data.
- Then we will study independent samples.

Outline

- 1 Introduction
- 2 Independent vs. Dependent Samples
- 3 Dependent Samples (Paired Data)**
- 4 Assignment

Dependent Samples (Paired Data)

- Let the pairs be denoted (x_1, x_2) .
- Let $d = x_2 - x_1$.
- We will study the case where d has a normal distribution.
- Let μ_D denote the mean of this distribution and σ_D denote the standard deviation.

Hypothesis Tests Concerning μ_D

- The only null hypothesis for μ_D that we will consider is

$$H_0 : \mu_D = 0.$$

- We will consider any of the three alternatives

$$H_1 : \mu_D < 0.$$

$$H_1 : \mu_D > 0.$$

$$H_1 : \mu_D \neq 0.$$

Hypothesis Tests Concerning μ_D

- For large samples, the test statistic is

$$t = \frac{\bar{d} - 0}{s_D / \sqrt{n}}.$$

(or z)

- For small samples it is necessary that d have a normal distribution. Then the test statistic is

$$t = \frac{\bar{d} - 0}{s_D / \sqrt{n}}.$$

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

- Suppose that a group of 10 students take a math placement test.
- Let the variable x_1 represent their scores on that test.
- Then they are given an Algebra refresher course and they retake the placement test.
- Let the variable x_2 represent their scores on the retest.

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

- The following table shows the results

Student	1st Score (x_1)	2nd Score (x_2)	Difference (d)
1	83	81	
2	62	63	
3	80	76	
4	73	80	
5	68	78	
6	67	71	
7	68	69	
8	69	78	
9	80	88	
10	83	79	

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

- The following table shows the results

Student	1st Score (x_1)	2nd Score (x_2)	Difference (d)
1	83	81	-2
2	62	63	1
3	80	76	-4
4	73	80	7
5	68	78	10
6	67	71	4
7	68	69	1
8	69	78	9
9	80	88	8
10	83	79	-4

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

- Test the hypothesis, at the 10% level, that the refresher course improved their grades on the placement test.

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(1) Let x_1 be the first test score, let x_2 be the second test score, and let $d = x_2 - x_1$.

Then the hypotheses are

$$H_0 : \mu_D = 0.$$

$$H_1 : \mu_D > 0.$$

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(1) Let x_1 be the first test score, let x_2 be the second test score, and let $d = x_2 - x_1$.

Then the hypotheses are

$$H_0 : \mu_D = 0.$$

$$H_1 : \mu_D > 0.$$

(2) $\alpha = 0.10$.

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

- (1) Let x_1 be the first test score, let x_2 be the second test score, and let $d = x_2 - x_1$.

Then the hypotheses are

$$H_0 : \mu_D = 0.$$

$$H_1 : \mu_D > 0.$$

- (2) $\alpha = 0.10$.

- (3) Let $t = \frac{\bar{d} - 0}{s_D / \sqrt{n}}$.

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(4) Compute the value of the test statistic.

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(4) Compute the value of the test statistic.

- Enter the x_1 values into L_1 and the x_2 values into L_2 .

Example (Hypothesis Tests Concerning μ_D)

(4) Compute the value of the test statistic.

- Enter the x_1 values into L_1 and the x_2 values into L_2 .
- Evaluate the difference $L_2 - L_1$ and store it in L_3 .

Example (Hypothesis Tests Concerning μ_D)

(4) Compute the value of the test statistic.

- Enter the x_1 values into L_1 and the x_2 values into L_2 .
- Evaluate the difference $L_2 - L_1$ and store it in L_3 .
- Use 1-Var Stats L_3 to get \bar{d} and s_D .

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(4) Compute the value of the test statistic.

- Enter the x_1 values into L_1 and the x_2 values into L_2 .
- Evaluate the difference $L_2 - L_1$ and store it in L_3 .
- Use 1-Var Stats L_3 to get \bar{d} and s_D .
- We find that $\bar{d} = 3$ and $s_D = 5.354$.

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(4) Compute the value of the test statistic.

- Enter the x_1 values into L_1 and the x_2 values into L_2 .
- Evaluate the difference $L_2 - L_1$ and store it in L_3 .
- Use 1-Var Stats L_3 to get \bar{d} and s_D .
- We find that $\bar{d} = 3$ and $s_D = 5.354$.
- Then

$$t = \frac{3}{5.354/\sqrt{10}} = \frac{3}{1.693} = 1.772.$$

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(5) $p\text{-value} = \text{tcdf}(1.772, E99, 9) = 0.0551.$

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(5) $p\text{-value} = \text{tcdf}(1.772, E99, 9) = 0.0551.$

(6) Reject $H_0.$

Hypothesis Tests Concerning μ_D

Example (Hypothesis Tests Concerning μ_D)

(5) $p\text{-value} = \text{tcdf}(1.772, E99, 9) = 0.0551.$

(6) Reject H_0 .

(7) Students scores on the placement test are higher after taking the Algebra refresher course.

Outline

- 1 Introduction
- 2 Independent vs. Dependent Samples
- 3 Dependent Samples (Paired Data)
- 4 Assignment**

Assignment

Homework

- Read Sections 11.1, 11.2, 11.3, pages 669 - 688.
- Let's Do It! 11.1, 11.2, 11.3.
- Exercises 1 - 8, page 676.
- Exercises 9 -14, page 689.