

Lecture 17: Paired Data and Difference of Two Means

Chapter 5.2, 5.1

Goals for Today

- ▶ Difference of means
- ▶ Note on Practical vs Statistical Significance
- ▶ Paired differences of means

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?
2. Are the means μ_1 and μ_2 of two groups different?

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?
2. Are the means μ_1 and μ_2 of two groups different?
3. What is the mean paired difference μ_{diff} ?

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?
2. Are the means μ_1 and μ_2 of two groups different?
3. What is the mean paired difference μ_{diff} ?
4. What is the proportion p of “successes”?

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?
2. Are the means μ_1 and μ_2 of two groups different?
3. What is the mean paired difference μ_{diff} ?
4. What is the proportion p of “successes”?
5. Are the proportions of “successes” p_1 and p_2 of two groups different?

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?
2. Are the means μ_1 and μ_2 of two groups different?
3. What is the mean paired difference μ_{diff} ?
4. What is the proportion p of “successes”?
5. Are the proportions of “successes” p_1 and p_2 of two groups different?
6. Are the means μ_1, \dots, μ_k of k groups different?

6 Types of Questions

Here are the 6 broad types of questions about **population parameters** we'll be answering with statistical methods: confidence intervals and hypothesis tests

1. What is the mean value μ ?
2. Are the means μ_1 and μ_2 of two groups different?
3. What is the mean paired difference μ_{diff} ?
4. What is the proportion p of “successes”?
5. Are the proportions of “successes” p_1 and p_2 of two groups different?
6. Are the means μ_1, \dots, μ_k of k groups different?

Today we look at 3 and 2.

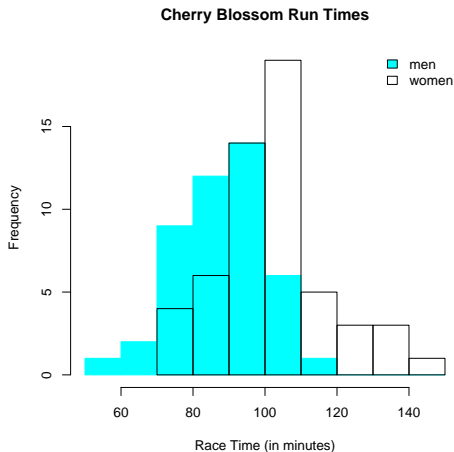
General Outline

Chapter 5.2: Are Two Means μ_1 & μ_2 Different?

We randomly sample 45 men (of 7192) and 55 women (of 9732) runners in the 2012 Cherry Blossom Run. Did men run faster than women?

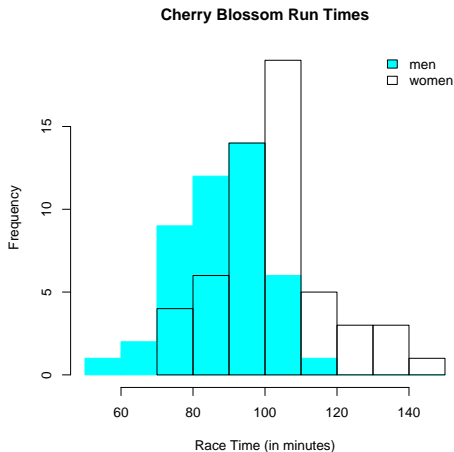
Chapter 5.2: Are Two Means μ_1 & μ_2 Different?

We randomly sample 45 men (of 7192) and 55 women (of 9732) runners in the 2012 Cherry Blossom Run. Did men run faster than women?



Chapter 5.2: Are Two Means μ_1 & μ_2 Different?

We randomly sample 45 men (of 7192) and 55 women (of 9732) runners in the 2012 Cherry Blossom Run. Did men run faster than women?



	men	women
\bar{x}	87.65	102.13
s	12.5	15.2
n	45	55

Practical vs Statistical Significance

When rejecting H_0 , we call this a statistically significant result. But statistically significant results aren't always practically significant.

Practical vs Statistical Significance

When rejecting H_0 , we call this a statistically significant result. But statistically significant results aren't always practically significant.

Say for very large n_M & n_F we observe $\bar{x}_M = 87.65$ and $\bar{x}_F = 87.651$ and reject H_0 .

Practical vs Statistical Significance

When rejecting H_0 , we call this a statistically significant result. But statistically significant results aren't always practically significant.

Say for very large n_M & n_F we observe $\bar{x}_M = 87.65$ and $\bar{x}_F = 87.651$ and reject H_0 .

The point estimate of the difference $\bar{x}_M - \bar{x}_F = 0.001$. Near negligible!

Practical vs Statistical Significance

When rejecting H_0 , we call this a statistically significant result. But statistically significant results aren't always practically significant.

Say for very large n_M & n_F we observe $\bar{x}_M = 87.65$ and $\bar{x}_F = 87.651$ and reject H_0 .

The point estimate of the difference $\bar{x}_M - \bar{x}_F = 0.001$. Near negligible!

The 95% CI might be:

$$[0.0005, 0.0015]$$

Practical vs Statistical Significance

Chapter 5.1: Paired Data

Chapter 5.1: Paired Data

Examples:

- ▶ Cholesterol levels before and after some intervention for the same person

Chapter 5.1: Paired Data

Examples:

- ▶ Cholesterol levels before and after some intervention for the same person
- ▶ Disease rates amongst pairs of twins

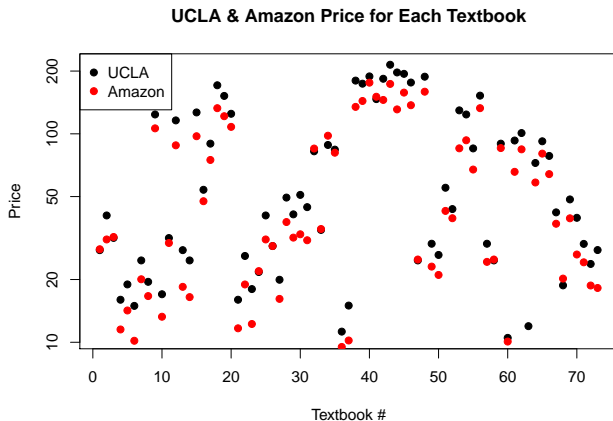
Chapter 5.1: Paired Data

Examples:

- ▶ Cholesterol levels before and after some intervention for the same person
- ▶ Disease rates amongst pairs of twins
- ▶ In the text: price of the same textbook at the UCLA bookstore vs Amazon

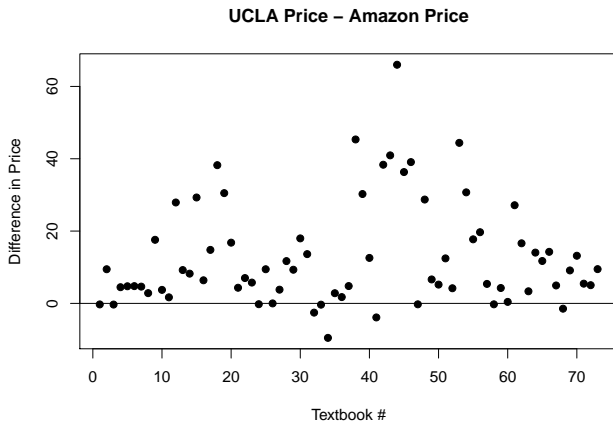
Paired Differences

The methodology for paired data remains the same, except our **observations** are the difference in pairs. Example, for the UCLA Bookstore vs Amazon book price example in the text



Paired Differences

The methodology for paired data remains the same, except our **observations** are the difference in pairs. Example, for the UCLA Bookstore vs Amazon book price example in the text



Paired Differences