# Hypothesis Testing

## Two Sample Means

# Moving from One Sample to Two

In this section, we are going to analyze the difference of means between two samples of data.

# Moving from One Sample to Two

In this section, we are going to analyze the difference of means between two samples of data.

In particular, we will examine whether there is a difference in means that is statistically significant, assuming there is no difference in means of each sample's population.

# Moving from One Sample to Two

In this section, we are going to analyze the difference of means between two samples of data.

In particular, we will examine whether there is a difference in means that is statistically significant, assuming there is no difference in means of each sample's population.

To put it mathematically, if we denote our first sample as $X$, and our second sample as $Y$, then we will want to know if

$$\mu_X - \mu_Y = 0$$

# Moving from One Sample to Two

In this section, we are going to analyze the difference of means between two samples of data.

In particular, we will examine whether there is a difference in means that is statistically significant, assuming there is no difference in means of each sample's population.

To put it mathematically, if we denote our first sample as $X$, and our second sample as $Y$, then we will want to know if

$$\mu_X - \mu_Y = 0 \quad \text{or} \quad \mu_X = \mu_Y$$

# Moving from One Sample to Two

In this section, we are going to analyze the difference of means between two samples of data.

In particular, we will examine whether there is a difference in means that is statistically significant, assuming there is no difference in means of each sample's population.

To put it mathematically, if we denote our first sample as $X$, and our second sample as $Y$, then we will want to know if

$$\mu_X - \mu_Y = 0 \quad \text{or} \quad \mu_X = \mu_Y$$

Thus, our null hypotheses will be $\mu_X = \mu_Y$.

# Objectives

1. Perform hypothesis test on the mean for two dependent samples
2. Perform hypothesis test on the mean for two independent samples

# Dependent Samples

Recall from probability that two events are **dependent** if the chance of the second event happening is affected by the first event happening.

# Dependent Samples

Recall from probability that two events are **dependent** if the chance of the second event happening is affected by the first event happening.

For this part of the section, we will look at examples with a "before-and-after" theme; much like with experiments.

# Dependent Samples

Recall from probability that two events are **dependent** if the chance of the second event happening is affected by the first event happening.

For this part of the section, we will look at examples with a "before-and-after" theme; much like with experiments.

This part is known as a **paired $t$ test**.

# Nuts and Bolts of a Paired $t$ Test

We will assume the distribution in the differences of sample $X$ and sample $Y$ are normal.

# Nuts and Bolts of a Paired $t$ Test

We will assume the distribution in the differences of sample $X$ and sample $Y$ are normal.

The test statistic is

$$t = \frac{\overline{d} - 0}{s_d / \sqrt{n}}$$

# Nuts and Bolts of a Paired $t$ Test

We will assume the distribution in the differences of sample $X$ and sample $Y$ are normal.

The test statistic is

$$t = \frac{\overline{d} - 0}{s_d / \sqrt{n}} = \frac{\overline{d}}{s_d / \sqrt{n}}$$

# Nuts and Bolts of a Paired $t$ Test

We will assume the distribution in the differences of sample $X$ and sample $Y$ are normal.

The test statistic is

$$t = \frac{\overline{d} - 0}{s_d/\sqrt{n}} = \frac{\overline{d}}{s_d/\sqrt{n}}$$

$\overline{d}$ represents the mean differences between the samples.

$$\overline{d} = \overline{X - Y}$$

# Nuts and Bolts of a Paired $t$ Test

We will assume the distribution in the differences of sample $X$ and sample $Y$ are normal.

The test statistic is

$$t = \frac{\overline{d} - 0}{s_d / \sqrt{n}} = \frac{\overline{d}}{s_d / \sqrt{n}}$$

$\overline{d}$ represents the mean differences between the samples.

$$\overline{d} = \overline{X - Y}$$

The population difference, $\mu_d$ will be 0 for hypothesis testing.

# Nuts and Bolts of a Paired $t$ Test

We will assume the distribution in the differences of sample $X$ and sample $Y$ are normal.

The test statistic is

$$t = \frac{\overline{d} - 0}{s_d/\sqrt{n}} = \frac{\overline{d}}{s_d/\sqrt{n}}$$

$\overline{d}$ represents the mean differences between the samples.

$$\overline{d} = \overline{X - Y}$$

The population difference, $\mu_d$ will be 0 for hypothesis testing.

$s$ represents the sample standard deviation of the differences, and there are $n - 1$ degrees of freedom.

# Example 1

A medication is given to patients in an attempt to lower their LDL cholesterol. The tables below list the levels. At the 5% significance level, test the claim that the medicine is effective in lowering LDL cholesterol.

| **Before** | **After** |
|---|---|
| 95 | 91 |
| 109 | 107 |
| 127 | 129 |
| 131 | 125 |
| 117 | 110 |
| 135 | 120 |
| 103 | 97 |
| 98 | 101 |
| 111 | 107 |

# Example 1

$$H_0 : \mu_{\text{Before}} = \mu_{\text{After}}$$

$$H_A : \mu_{\text{Before}} < \mu_{\text{After}}$$

# Example 1

$H_0 : \mu_{\text{Before}} = \mu_{\text{After}}$

$H_A : \mu_{\text{Before}} < \mu_{\text{After}}$

- $t = 2.446$

# Example 1

$H_0 : \mu_{\text{Before}} = \mu_{\text{After}}$

$H_A : \mu_{\text{Before}} < \mu_{\text{After}}$

- $t = 2.446$ (critical value $= 1.96$)

# Example 1

$H_0 : \mu_{\text{Before}} = \mu_{\text{After}}$

$H_A : \mu_{\text{Before}} < \mu_{\text{After}}$

- $t = 2.446$ (critical value $= 1.96$)
- $p = 0.0201$

# Example 1

$H_0 : \mu_{\text{Before}} = \mu_{\text{After}}$

$H_A : \mu_{\text{Before}} < \mu_{\text{After}}$

- $t = 2.446 \quad$ (critical value $= 1.96$)
- $p = 0.0201 \quad (\alpha = 0.05)$

# Example 1

$H_0$ : $\mu_{\text{Before}} = \mu_{\text{After}}$

$H_A$ : $\mu_{\text{Before}} < \mu_{\text{After}}$

- $t = 2.446 \quad$ (critical value $= 1.96$)
- $p = 0.0201 \quad (\alpha = 0.05)$
- 95% confidence interval: (0.2478, 8.4189)

# Example 1

$H_0 : \mu_{\text{Before}} = \mu_{\text{After}}$

$H_A : \mu_{\text{Before}} < \mu_{\text{After}}$

- $t = 2.446$ (critical value $= 1.96$)
- $p = 0.0201$ $(\alpha = 0.05)$
- 95% confidence interval: (0.2478, 8.4189) does not contain $\mu_d = 0$

# Example 1

Reject the null hypothesis.

# Example 1

Reject the null hypothesis.

> *At the 5% significance level we reject the null hypothesis that there is no difference in LDL cholesterol levels and conclude that our sample suggests the medication may be effective in lowering LDL cholesterol levels.*

# Objectives

1. Perform hypothesis test on the mean for two dependent samples

2. Perform hypothesis test on the mean for two independent samples

# Independent Samples

For this part, our samples will be *independent* of one another.

# Independent Samples

For this part, our samples will be *independent* of one another.

In other words, there is no "before-and-after" relationship between our samples, and our samples don't even have to be the same sizes.