Introduction to Statistics

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

November 10, 2019

Topics

1. Paired Samples Test for μ_1 and μ_2

Objectives

Objectives:

- Distinguish between paired samples and independent samples.
- Carry out a paired-samples t test for two population means μ_1 and μ_2.

Paired Samples Test for μ_1 and μ_2 (10.5)

Paired Samples

- ***Paired samples*** arise when **two samples** are collected in **pairs**, whereby each individual from the first population is **matched** with one in the second population, and a random sample of **pairs** is selected.

Paired Samples Test for μ_1 and μ_2 (10.5)

Paired Samples

- ***Paired samples*** arise when **two samples** are collected in **pairs**, whereby each individual from the first population is **matched** with one in the second population, and a random sample of **pairs** is selected.

  Matching of individuals is done in such a way that **within each pair**, the two individuals have **similar characteristics**.

Exercise

Decide which of the following are **paired samples** and which are **independent samples**.

Exercise

Decide which of the following are **paired samples** and which are **independent samples**.

a) To compare prices at two grocery stores, a random sample of food items is selected from one store, and the **same** items selected from the other store.

Exercise

Decide which of the following are **paired samples** and which are **independent samples**.

a) To compare prices at two grocery stores, a random sample of food items is selected from one store, and the **same** items selected from the other store.

b) To compare lengths of comedy and drama movies, a random sample of comedies is selected and a **separate** random sample of dramas is selected.

- **Pairs samples** can also arise when **two measurements** are made **on each individual** in a study, one measurement under one condition and the other under a different condition.

Example

To decide if college students retain information better by reading the material in a **book** or by watching it on a **DVD**, a sample of $n = 7$ students was each was assigned one 19th-century novel to read and a different one to watch.

Example

To decide if college students retain information better by reading the material in a **book** or by watching it on a **DVD**, a sample of $n = 7$ students was each was assigned one 19th-century novel to read and a different one to watch.

Afterward, they were given a 100-point **quiz** on each novel.

These **two measurements** (quiz scores) for **each student** resulted in the following **paired samples**:

Student	Book Quiz Result	DVD Quiz Result
1	90	85
2	80	72
3	90	80
4	75	80
5	80	70
6	90	75
7	84	80

Paired Differences

- The **pairwise *differences*** are obtained by subtracting each value in the second sample from its paired value in the first sample.

Paired Differences

- The **pairwise** ***differences*** are obtained by subtracting each value in the second sample from its paired value in the first sample.

  This is illustrated in the next example.

Example

For the study to decide if college students retain information better by reading a book or by watching a DVD, the seven **paired differences** are shown in the right column below:

Student	Book Test Result	DVD Test Result	**Difference**
1	90	85	**5**
2	80	72	**8**
3	90	80	**10**
4	75	80	**-5**
5	80	70	**10**
6	90	75	**15**
7	84	80	**4**

- When the samples are ***paired***, we **shouldn't** use the two-mean t test for μ_1 and μ_2 (because that test is only appropriate for *independent* samples).

  Instead, use the ***paired samples t test*** described ahead.

- When the samples are ***paired***, we **shouldn't** use the two-mean t test for μ_1 and μ_2 (because that test is only appropriate for *independent* samples).

  Instead, use the ***paired samples t test*** described ahead.

  (The test can also be used to compare responses to **two treatments** in a ***paired*** **experiment**.)

- The test considers the **paired differences** to be a **single random sample** from a (hypothetical) ***"population of differences"***.

- The test considers the **paired differences** to be a **single random sample** from a (hypothetical) ***"population of differences"***.

  We'll use the sample of *differences* to decide if the population means μ_1 and μ_2 are different.

- The **null hypothesis** is:

Null Hypothesis:

$$H_0 : \ \mu_1 \ = \ \mu_2$$

- The **alternative hypothesis** will depend on what we're trying to **substantiate**:

Alternative Hypothesis: The alternative hypothesis will be one of

1. $H_a: \mu_1 > \mu_2$ (**one-sided, upper-tailed test**)
2. $H_a: \mu_1 < \mu_2$ (**one-sided, lower-tailed test**)
3. $H_a: \mu_1 \neq \mu_2$ (**two-sided, two-tailed test**)

depending on what we're trying to substantiate in our study.

- The **test statistic** for the ***paired samples*** t ***test*** is

Paired t **Test Statistic**:

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

where

$$\bar{d} = \text{The sample mean of the } n \textit{ differences}$$

and

$$s_d = \text{The sample standard deviation of the } n \textit{ differences}$$

When

$$H_0 : \mu_1 = \mu_2$$

is true, the **sampling distribution** of the test statistic t is a t distribution with $n - 1$ df.

When

$$H_0 : \mu_1 = \mu_2$$

is true, the **sampling distribution** of the test statistic t is a t distribution with $n - 1$ df.

(The *paired samples* t *test* is valid if either the sample of *differences* is from a **normal** population or the sample size n is **large** ($n \geq 30$).)

- **Intuition**: It can be shown that the **mean** of the ***"population of differences"***, denoted μ_d, is equal to the difference between the two population means, i.e.

$$\mu_d = \mu_1 - \mu_2.$$

Thus the **paired samples t test statistic** can be viewed as a **one-mean t test statistic** for

$$H_0 : \mu_d = 0$$

using the **sample of *differences***.

- **Intuition**: It can be shown that the **mean** of the ***"population of differences"***, denoted μ_d, is equal to the difference between the two population means, i.e.

$$\mu_d = \mu_1 - \mu_2.$$

  Thus the **paired samples t test statistic** can be viewed as a **one-mean t test statistic** for

$$H_0 : \mu_d = 0$$

  using the **sample of *differences***.

- t measures (approximately) how many **standard errors** $\bar{d}$ is **away from** the claimed value **zero**.

1. Values of t **close to zero** provide almost ***no* evidence against** the **null hypothesis** $H_0 : \mu_1 = \mu_2$.

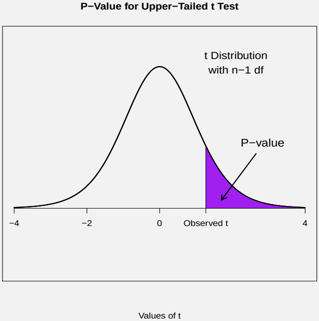
2. **Positive** values of t provide **evidence against the null hypothesis in favor of** $H_a : \mu_1 > \mu_2$.

3. **Negative** values of t provide **evidence against the null hypothesis in favor of** $H_a : \mu_1 < \mu_2$.

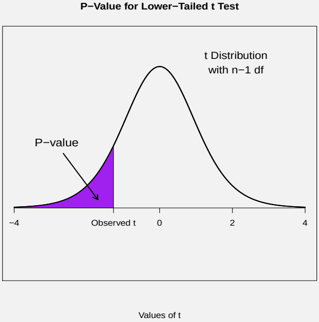
4. **Positive *and* negative** values of t provide **evidence against the null hypothesis in favor of** $H_a : \mu \neq \mu_0$.

- The ***p-value*** is the probability that just by chance (under H_0) we'd get a test statistic value as far from zero, in the direction predicted by H_a, as the observed value.

1. **P-value** = Area to the **right** of the observed t if the alternative hypothesis is $H_a : \mu > \mu_0$.

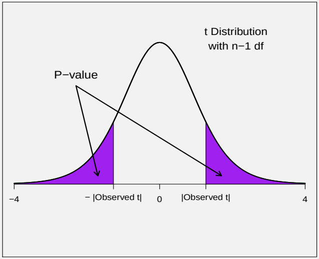


2. **P-value** = Area to the **left** of the observed t if the alternative hypothesis is $H_a : \mu < \mu_0$.



3. **P-value** = Area to the **left** of $-|t|$ **and right** of $|t|$ if the alternative hypothesis is $H_a : \mu \neq \mu_0$.

P−Value for Two−Tailed t Test



Values of t

- As always, after choosing a **level of significance** α, we **reject** H_0 if **p-value** $\leq \alpha$, otherwise fail reject H_0.

Example

A company wanted to know if attending a course on ”how to be a successful salesperson” can increase the average sales of its employees.

Example

A company wanted to know if attending a course on ”how to be a successful salesperson” can increase the average sales of its employees.

The company sent $n = 6$ of its salespersons to attend this course.

Example

A company wanted to know if attending a course on "how to be a successful salesperson" can increase the average sales of its employees.

The company sent $n = 6$ of its salespersons to attend this course.

The table on the next slide gives the one-week sales of these salespersons **before** and **after** they attended this course, along with the **differences** (after minus before).

Salesperson	Sales After the Course	Sales Before the Course	Difference
1	18	12	6
2	24	18	6
3	24	25	-1
4	14	9	5
5	19	14	5
6	20	16	-4

Salesperson	Sales After the Course	Sales Before the Course	Difference
1	18	12	6
2	24	18	6
3	24	25	-1
4	14	9	5
5	19	14	5
6	20	16	-4

The summary statistics for the **paired differences** shown above are

$$\bar{d} = 4.17 \quad \text{and} \quad s_d = 2.64$$

We'll carry out a **paired-samples** t **test** to **decide if the sales increase** as a result of attending the course.

We'll carry out a **paired-samples t test** to **decide if the sales increase** as a result of attending the course.

The **hypotheses** are

$$\begin{aligned} H_0 : \mu_1 &= \mu_2 \\ H_a : \mu_1 &> \mu_2 \end{aligned}$$

where μ_1 is the true (unknown) population mean sales **after** attending the course, and μ_2 is the true mean **before.**

The observed **test statistic** is

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

The observed **test statistic** is

$$\begin{aligned} t &= \frac{\bar{d} - 0}{s_d/\sqrt{n}} \\ &= \frac{4.17 - 0}{2.64/\sqrt{6}} \end{aligned}$$

The observed **test statistic** is

$$\begin{aligned} t &= \frac{\bar{d} - 0}{s_d/\sqrt{n}} \\ &= \frac{4.17 - 0}{2.64/\sqrt{6}} \\ &= \mathbf{3.87}. \end{aligned}$$

Thus the **sample mean *difference*** in sales, $\bar{d} = \mathbf{4.17}$, is about **3.87 standard errors above zero.**

The observed **test statistic** is

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}} = \frac{4.17 - 0}{2.64/\sqrt{6}} = \mathbf{3.87}.$$

Thus the **sample mean *difference*** in sales, $\bar{d} = 4.17$, is about **3.87 standard errors above zero.**

The **p-value** is the **probability** that we'd get a t value this far above zero (in either direction) **if** the **course had no effect** on sales.

From the **upper tail** of the t **distribution** with $n - 1 = 5$ **df**, to the **right** of **3.87**,

$$\textbf{p-value} \; = \; 0.006$$

(obtained from the **table** showing **areas to the right of** t**).**

From the **upper tail** of the t **distribution** with $n - 1 = 5$ **df**, to the **right** of **3.87**,

$$\textbf{p-value} \;=\; 0.006$$

(obtained from the **table** showing **areas to the right of** t).

Thus we'd get a result like the one we got **only 0.6%** of the time **if** the **course had no effect** on sales.

Using a **level of significance** $\boldsymbol{\alpha = 0.05}$, the **decision rule** is

Reject H_0 if p-value < 0.05.
Fail to reject H_0 if p-value ≥ 0.05.

Using a **level of significance** $\alpha = \mathbf{0.05}$, the **decision rule** is

Reject H_0 if p-value < 0.05.
Fail to reject H_0 if p-value ≥ 0.05.

Because $\mathbf{0.006 < 0.05}$, we **reject** $\boldsymbol{H_0}$.

Using a **level of significance** $\alpha = 0.05$, the **decision rule** is

Reject H_0 if p-value < 0.05.
Fail to reject H_0 if p-value ≥ 0.05.

Because $0.006 < 0.05$, we **reject** H_0.

There's **statistically significant evidence** that the course results in increased sales.

Using a **level of significance** $\alpha = 0.05$, the **decision rule** is

Reject H_0 if p-value < 0.05.
Fail to reject H_0 if p-value ≥ 0.05.

Because $\mathbf{0.006 < 0.05}$, we **reject** $\boldsymbol{H_0}$.

There's **statistically significant evidence** that the course results in increased sales.

The observed differences in sales (for the sample of six employed who attended the course) *cannot* be explained by chance variation (sampling error).

Exercise

Refer to the study to decide if college students retain information better by reading a book or by watching a DVD.

Exercise

Refer to the study to decide if college students retain information better by reading a book or by watching a DVD.

Using the $n = 7$ paired differences given in that example, the summary statistics for the **differences** are

$$\bar{d} = 6.71 \qquad \text{and} \qquad s_d = 6.32$$

Carry out the **paired-samples t test** to **decide if there's any difference** in a student's ability to retain information by reading a **book** and by watching a **DVD**. Use a level of significance $\alpha = 0.05$.