Two Sample T-Distribution

QMB 3200

09/24/2020

Paired Data or Dependent Means t-test

Paired Data

1. Use the sample as its own control
    1. Collect data
    2. Administer experiment/intervention
    3. Collect data again
    4. Determine whether the experiment changed the sample's mean score
2. Use a "control" sample
    1. Randomly assign participants to one of two conditions: (0) Control or (1) Experimental
    2. Conduct experiment
    3. Collect data on the two samples
    4. Compare their mean scores

Paired Data or Dependent Samples t-test

- A dependence exists in the data
  - Same subjects measured twice
  - Related subjects
    - Mother-Father evaluation of child's behavior

- The same dependent variable is measured twice on the same subjects *OR* on related subjects
  - Repeated Measures
    - Information at Time 1 conveys information at Time 2
  - Related Subjects
    - Information collected from one subject conveys information about the related subject

Examples of Potential Research Questions Requiring a Dependent Samples t-test

- Does training improve scores?

- Do mothers and fathers rate their child the same, or differently on a measure of externalizing behavior?

- Did cognitive abilities change from age 2 to age 3 for the same individuals?

Dependent Samples t-test

- In these situations, we want to know if there is a difference between (or change in) the scores
  - Difference between pre- and post-test scores
  - Difference between mother’s and father’s scores
  - Difference in scores collected at age 2 and 3 for the same sample of infants

Appropriate Statistical Test for a Dependent Samples t-test

- Based on the Difference Score: *D*
  - $D_{ki} = X_{ki2} - X_{ki1}$
- What's the expected value of the difference score *under the Null Hypothesis*
  - Training did not improve scores
  - Mothers and fathers do not differ in their ratings of externalizing behavior
  - Children's cognitive abilities did not develop from age 2 to 3
- Compare mean of the difference score to 0, given the standard error of the mean of the difference score

Appropriate Statistical Test for a Dependent Samples t test

- Appropriate *t*-value

$$t_{\overline{D}} = \frac{\overline{D} - 0}{\frac{s_D}{\sqrt{n}}}$$

$\overline{D}$ – is the mean of the difference scores

0 – is the expected mean under H_0 (e.g., $\mu_{t=1} = \mu_{t=2}$)

s_D – is the standard deviation of the difference scores

n – is the sample size

Example

- Dr. Bates wanted to examine the effect of a cognitive training paradigm. Nine students participated in the study and their scores on a fluid intelligence test were recorded before and after the cognitive training paradigm.

Setting up the Hypotheses

- H_0: The training had no effect
  - $\mu_{t=1} = \mu_{t=2}$ OR $\mu_D = 0$
- H_1:The training had an effect (changed the mean)
  - $\mu_{t=1} \neq \mu_{t=2}$ OR $\mu_D \neq 0$

Cognitive Scores

	Time 1	Time 2	Difference
$i = 1$	13	17	
$i = 2$	12	18	
$i = 3$	14	17	
$i = 4$	14	16	
$i = 5$	12	21	
$i = 6$	13	18	
$i = 7$	18	20	
$i = 8$	14	16	
$i = 9$	16	18	
Mean (M)	14.00	17.89	
Standard Deviation (s)	1.94	1.69	

Cognitive Scores

	Time 1	Time 2	Difference
$i = 1$	13	17	+4
$i = 2$	12	18	
$i = 3$	14	17	
$i = 4$	14	16	
$i = 5$	12	21	
$i = 6$	13	18	
$i = 7$	18	20	
$i = 8$	14	16	
$i = 9$	16	18	
Mean (M)	14.00	17.89	
Standard Deviation (s)	1.94	1.69	

Cognitive Scores

	Time 1	Time 2	Difference
$i = 1$	13	17	+4
$i = 2$	12	18	+6
$i = 3$	14	17	+3
$i = 4$	14	16	+2
$i = 5$	12	21	+9
$i = 6$	13	18	+5
$i = 7$	18	20	+2
$i = 8$	14	16	+2
$i = 9$	16	18	+2
Mean (M)	14.00	17.89	
Standard Deviation (s)	1.94	1.69	

Calculate Mean & Standard Deviation of Difference Score

- Mean

$$M = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{35}{9} = 3.89$$

- Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{N} (X_i - M)^2}{N-1}} = \sqrt{\frac{46.89}{8}} = \sqrt{5.86} = 2.42$$

Cognitive Scores

	Time 1	Time 2	Difference
$i = 1$	13	17	+4
$i = 2$	12	18	+6
$i = 3$	14	17	+3
$i = 4$	14	16	+2
$i = 5$	12	21	+9
$i = 6$	13	18	+5
$i = 7$	18	20	+2
$i = 8$	14	16	+2
$i = 9$	16	18	+2
Mean (*M*)	14.00	17.89	3.89
Standard Deviation (*s*)	1.94	1.69	2.42
S.E. of the Mean (*se*)	0.65	0.56	

Standard Error of the Mean Difference

- Standard Error of the Mean Difference

$$s_{\overline{D}} = \frac{s_D}{\sqrt{n}} = \frac{2.42}{\sqrt{9}} = .81$$

Cognitive Scores

	Time 1	Time 2	Difference
$i = 1$	13	17	+4
$i = 2$	12	18	+6
$i = 3$	14	17	+3
$i = 4$	14	16	+2
$i = 5$	12	21	+9
$i = 6$	13	18	+5
$i = 7$	18	20	+2
$i = 8$	14	16	+2
$i = 9$	16	18	+2
Mean (M)	14.00	17.89	3.89
Standard Deviation (s)	1.94	1.69	2.42
S.E. of the Mean (se)	0.65	0.56	0.81

df for Dependent Samples t-test

- For repeated measures (e.g., effect of training)
  - *df* = sample size minus 1 (*N* – *1*)

- For a matched sample (e.g., Mother & Father rating of child's externalizing behavior)
  - *df* is the <u>number of pairs</u> minus 1

t-value for example

- *t*-value (Observed)

$$t = \frac{M_{\overline{D}} - 0}{\frac{s_D}{\sqrt{n}}} = \frac{M_{\overline{D}}}{\frac{s_D}{\sqrt{n}}} = \frac{3.89}{.81} = 4.802$$

- *Critical t*(*8*) = 2.306
- Decision:
  - Reject the Null Hypothesis
- Conclusion:
  - The training improved cognitive scores

95% Confidence Interval of the Mean Difference

$$CI_{0.95} = \overline{D} \pm t_{0.025} \cdot s_{\overline{D}}$$

$$= 3.89 \pm 2.306 \cdot (.81)$$

$$= 3.89 \pm 1.87$$

$$CI_{0.95} = \left[1.91, 5.76\right]$$

Second Example

Hoaglin, Mosteller, and Tukey (1983) present data on blood levels of beta-endorphin as a function of stress. They took beta-endorphin levels for 19 patients 12 hours before surgery, and again 10 minutes before surgery.

What can they conclude regarding changes in the blood levels of beta-endorphin?

Data

ID	12 hours	10 minutes	Difference
1	10	6.5	
2	6.5	14.0	
3	8.0	13.5	
4	12.0	18.0	
5	5.0	14.5	
6	11.5	9.0	
7	5.0	18.0	
8	3.5	42.0	
9	7.5	7.5	
10	5.8	6.0	

ID	12 hours	10 minutes	Difference
11	4.7	25.0	
12	8.0	12.0	
13	7.0	52.0	
14	17.0	20.0	
15	8.8	16.0	
16	17.0	15.0	
17	15.0	11.5	
18	4.4	2.5	
19	2.0	2.0	

Data

ID	12 hours	10 minutes	Difference
1	10	6.5	3.5
2	6.5	14.0	-7.5
3	8.0	13.5	-5.5
4	12.0	18.0	-6.0
5	5.0	14.5	-9.5
6	11.5	9.0	2.5
7	5.0	18.0	-13
8	3.5	42.0	-38.5
9	7.5	7.5	0
10	5.8	6.0	-0.2

ID	12 hours	10 minutes	Difference
11	4.7	25.0	-20.3
12	8.0	12.0	-4
13	7.0	52.0	-45
14	17.0	20.0	-3.0
15	8.8	16.0	-7.2
16	17.0	15.0	2.0
17	15.0	11.5	3.5
18	4.4	2.5	1.9
19	2.0	2.0	0

Example

- Hypotheses:
  - H_0: $\mu_D = 0$
  - H_1: $\mu_D \neq 0$

- Data:
  - Difference Scores (t_2 - t_1)
    - $n = 19$, *Mean* $= 7.70$, *St. Dev.* $= 13.52$

Second Example

- Calculate *t*-value for the mean difference
  - Observed mean difference divided by the standard error.

$$t = \frac{\overline{X}_D}{\frac{s_D}{\sqrt{n}}} = \frac{7.70}{\frac{13.52}{\sqrt{19}}} = \frac{7.70}{3.10} = 2.48$$

Second Example

- Determine critical value for the *t*-distribution given our *degrees of freedom*
  - *df = n - 1*
  - *df = 18*
- Critical *t-value = 2.101*
- Decision
  - $t_{OBS} > 2.101$
  - *Reject the Null Hypothesis*
- Conclusion
  - *The level of beta-endorphin significantly increased 12 hours before to 10 minutes before surgery. Thus, it appears that the stress of being closer to surgery raised beta-endorphin blood levels.*

Independent Samples t-test or Difference of two means

Independent Samples t-test

- Two separate groups, no interdependence
  - Experimental and Control
  - Intact Groups
    - Males & Females
    - Upperclassmen vs. Lowerclassmen
- Measured on the same dependent variable
- Independence of observations

Issue

- As we've seen with the Dependent Means *t*-test, we are interested in the *Sampling Distribution* of the *Mean Difference*.
- However, if our sample is not dependent, there's no way to calculate difference scores at the ***individual*** level.
- Therefore, everything must be done at the ***group*** level.
- We know the mean of the *Sampling Distribution* of the mean difference is ***zero***, but we don't know the standard deviation of this sampling distribution (just the standard deviation for each group)

Variance Calculation

- Variance Sum Law
  - The variance of a sum or difference of two *independent* variables is equal to the sum of their variances

- Thus

$$\sigma^2_{\overline{X}_1+\overline{X}_2} = \sigma^2_{\overline{X}_1-\overline{X}_2} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Independent Samples t-test

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \cancel{\left(\mu_1 - \mu_2\right)}}{s_{\left(\overline{X}_1 - \overline{X}_2\right)}}$$

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Independent Samples t-test

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$df = n_1 + n_2 - 2$

Independent Samples t-test

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

This works well if the sample sizes are equivalent

What if the sample sizes are not equivalent?

Unequal sample sizes

- When the sample size for the two independent samples/groups is not equal
- Take the *weighted* average because the variance estimate will be more precise as sample size increases

Pooled Variance

- If the sample sizes are not equivalent, then we need to pool their variances – weighted by sample size

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Independent Samples t-test Unequal Sample Sizes (General Formula)

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{s_{\overline{X}_1 - \overline{X}_2}} = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$df = n_1 + n_2 - 2$

$n_1 + n_2$ independent pieces of information

2 mean estimates leaves us with $(n_1 + n_2 - 2)$ degrees freedom

Example

Aronson and colleagues (1998) studied the effects of stereotype threat. In this study, 23 Caucasian students were administered a difficult math test. Twelve of the students were told that Asian students typically did better than other students in math tests, and that the purpose of the exam was to help the experimenter understand why this difference exists (Threat Condition). Eleven students were simply asked to take the test (Control Condition).

Example

- Hypotheses:
  - H_0: $\mu_1 = \mu_2$
  - H_1: $\mu_1 \neq \mu_2$
- Data:
  - Control Condition
    - $n = 11$, *Mean* $= 9.64$, *St. Dev.* $= 3.17$
  - Threat Condition
    - $n = 12$, *Mean* $= 6.58$, *St. Dev.* $= 3.03$

Sampling Distribution of the Difference between Independent Sample Means

- Mean of Sampling Distribution
  - $\mu = 0$
- Variance of Sampling Distribution
  - Calculate Pooled Variance

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(11-1)3.17^2 + (12-1)3.03^2}{11+12-2}$$

$$= \frac{(10)10.0489 + (11)9.1809}{21}$$

$$= \frac{201.4789}{21} \qquad = 9.5942$$

Standard Error of the Mean Difference

$$se_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}} = \sqrt{\frac{9.5942}{11}+\frac{9.5942}{12}}$$

$$= \sqrt{1.6717}$$

$$= 1.2929$$

Example

- Calculate *t*-value for the mean difference
  - Observed mean difference divided by the standard error.

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{se_{\overline{X}_1 - \overline{X}_2}} = \frac{(9.64 - 6.58)}{1.2929}$$

$$= \frac{3.06}{1.2929}$$

$$= 2.37$$

Example

- Determine critical value for the *t*-distribution given our *degrees of freedom*
  - $df = n_1 + n_2 - 2$
  - $df = 21$
- Critical *t-value = 2.080*
- Decision
  - $t_{OBS} > 2.080$
  - *Reject the Null Hypothesis*
- Conclusion
  - *Subjects in the threat condition performed significantly worse than subjects in the control condition.*

95% Confidence Limits Around $\mu_1 - \mu_2$

- Estimated from sample mean differences

$$CI_{0.95} = (\overline{X}_1 - \overline{X}_2) \pm t_{critical} \cdot s_{\left(\overline{X}_1 - \overline{X}_2\right)}$$

Additional Example

Much has been made of the concept of experimenter bias, which refers to the fact that even the most conscientious experimenters tend to collect data that come out in the desired direction (they see what they want to see). Suppose we use students as experimenters. All experimenters are told that subjects will given caffeine before the experiment, but on-half of the experimenters are told that we expect caffeine to lead to good performance and one half are told that we expect it to lead to poor performance. The dependent variable is the number of simple arithmetic problems the subjects can solve in 2 minutes. The data obtained are as follows:

Expectation Good: n = 9, M = 18.78, s = 3.93

Expectation Bad: n = 8, M = 17.63, s = 4.17

What can you conclude?

Calculate the 95% confidence limits on $\mu_1 - \mu_2$.

Example

- Hypotheses:
  - H_0: $\mu_1 = \mu_2$
  - H_1: $\mu_1 \neq \mu_2$

- Data:
  - Good Condition
    - $n = 9$, $Mean = 18.78$, $St.\ Dev. = 3.93$
  - Bad Condition
    - $n = 8$, $Mean = 17.63$, $St.\ Dev. = 4.17$

Sampling Distribution of the Difference between Independent Sample Means

- Mean
  - $\mu = 0$
- Variance
  - Calculate Pooled Variance

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(9-1)3.93^2 + (8-1)4.17^2}{9+8-2}$$

$$= \frac{(8)15.4449 + (7)17.3889}{15}$$

$$= \frac{245.2815}{15} \qquad = 16.3521$$

Standard Error of the Mean Difference

$$se_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{16.3521}{9} + \frac{16.3521}{8}}$$

$$= \sqrt{3.8609}$$

$$= 1.9649$$

Example

- Calculate *t*-value for the mean difference
  - Observed mean difference divided by the standard error.

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{se_{\overline{X}_1 - \overline{X}_2}} = \frac{(18.78 - 17.63)}{1.9649}$$

$$= \frac{1.15}{1.9649}$$

$$= .585$$

Example

- Determine critical value for the *t*-distribution given our *degrees of freedom*
  - $df = n_1 + n_2 - 2$
  - $df = 15$
- Critical *t-value = 2.131*
- Decision
  - $t_{OBS} < 2.131$
  - *Retain the Null Hypothesis*
- Conclusion
  - *Subjects in the "Good" condition scored similarly to subjects in the "Bad" condition.*

Calculate Confidence Limits

- Calculate the 95% confidence limits on μ_1 - μ_2 for the data in previous example

$$CI_{0.95} = (\overline{X}_1 - \overline{X}_2) \pm t_{critical} \cdot s_{(\overline{X}_1 - \overline{X}_2)}$$

$$CI_{0.95} = (1.15) \pm 2.131 \cdot 1.9649$$

$$CI_{0.95} = [5.34, -3.04]$$

More Examples

- The following are the average weekly losses of worker-hours due to accidents in 10 industrial plants before and after a certain safety program was put into operation

Before	45	73	46	124	33	57	83	34	26	17
After	36	60	44	119	35	51	77	29	24	11

- Use the 0.05 level of significance to test whether the safety program is effective

More Examples

- The following are the average weekly losses of worker-hours due to accidents in 10 industrial plants before and after a certain safety program was put into operation

Before	45	73	46	124	33	57	83	34	26	17
After	36	60	44	119	35	51	77	29	24	11

- Use the 0.05 level of significance to test whether the safety program is effective

More Examples

- The dynamic modulus of concrete is obtained for two different concrete mixes. For the first mix, n_1 = 33, $\overline{x_1}$ = 115.1 and s_1 = 0.47 psi. for the second mix, n_2 = 31, $\overline{x_2}$ = 114.6 and s_2 = 0.38 psi. Test with 0.05 level of significance the null hypothesis of equality of mean dynamic modulus versus the two-sided alternative.

t-values (in general)

- *t*-value is equal to the statistic divided by the standard error of the statistic
  - Calculating the standard error depends on whether we have independent or dependent samples
- With the exception of the one-sample tests, we want to know if the statistic is significantly different from 0
- If the confidence interval of the statistic includes zero, then we say the statistic is not significantly different from zero.

Review

- Dependent Samples *t*-test
  - There is a logical connection between the data
    - e.g., repeated measures data
  - Calculate difference scores, mean of difference scores, and standard deviation of difference scores
  - Conduct a One-sample t-test with the information obtained from the difference scores with 0 as the population mean

Review

- Independent Samples *t*-test
  - Most common version of the *t*-test
  - Calculate Mean, Standard Deviation for each independent sample
  - Calculate the pooled variance, use pooled variance to calculate the *appropriate* standard error
  - Divide mean difference by the standard error and determine if value is greater than the critical value given the degrees of freedom

Review

- Comparing a sample mean to a population and the population mean and standard deviation are **known**
  - z-test
- Comparing a sample mean to a population and the population mean is **known**, but the standard deviation is **unknown**
  - *t*-test
- Confidence intervals are used to give a sense of how uncertain we are about the population mean. Confidence interval will shrink as *n* increases (holding everything else constant)