# Stat 250 Gunderson Lecture Notes
## 8: Learning about a Population Mean Difference

### Part 1: Distribution for a Sample Mean of Paired Differences

## The Paired Data Scenario

An important special case of a single mean of a population occurs when two quantitative variables are **collected in pairs**, and we desire information about the **difference** between the two variables.  Here are some ways that paired data can occur:

- Each person or unit is measured twice. The two measurements of the same characteristic or trait are made under different conditions. An example is measuring a quantitative response both before and after treatment.

- Similar individuals or units are paired prior to an experiment.   During the experiment, each member of a pair receives a different treatment.   The same quantitative response variable is measured for all individuals.

For paired data designs, it is the ***differences*** that we are interested in examining. By focusing on the differences we again have just one sample of observations (the differences).  Sometimes you may see a "*d*" in the subscript of the mean to represent the **mean of the population of differences**: $\mu_d$; and the data may be represented generically as: $d_1, d_2, ..., d_n$.

## Sampling Distribution for the Sample Mean of Paired Differences

The sampling distribution results for the mean of paired differences is really the same as that for a regular sample mean.  Since the measurements are differences, the sample mean of the data, $\bar{x}$, is just written as $\bar{d}$, to emphasize that this is a paired design.

### Freshmen Weight

A study was conducted to learn about the average weight gain in the first year of college for students.  A sample of 60 students resulted in an average weight gain of 4.2 pounds (over the first 12 weeks of college).

Population =

Parameter =

Sample =

Statistic =

Can anyone say how close this observed sample mean difference $\bar{d}$ of 4.2 pounds is to the true population mean difference $\mu_d$? _____   If we were to take another random sample of the same size, would we get the same value for the sample mean difference? _____. So what are the possible values for the sample mean difference $\bar{d}$ if we took many random samples of the same size from this population? What would the distribution of the possible $\bar{d}$ values look like?   What can we say about the **distribution of the sample mean difference**?

## Distribution of the Sample Mean Difference – Main Results

Let $\mu_d$ = mean of the differences in the population of interest.
Let $\sigma_d$ = standard deviation for the differences in the population of interest.
Let $\bar{d}$ = the sample mean of the differences for a random sample of size $n$.

- If the population of differences is normal (bell-shaped), and a random sample of any size is obtained, then the distribution of the sample mean difference $\bar{d}$ is also normal, with a mean of $\mu_d$ and a standard deviation of $s.d.(\bar{d}) = \dfrac{\sigma_d}{\sqrt{n}}$.

- If the population of differences is not normal (bell-shaped), but a *large* random sample of size $n$ is obtained, then the distribution of the sample mean difference $\bar{d}$ is *approximately* normal, with a mean of $\mu_d$ and a standard deviation of $s.d.(\bar{d}) = \dfrac{\sigma_d}{\sqrt{n}}$.

**Notes:**
(1) An arbitrary level for what is 'large' enough has been 30. However, if any of the differences are extreme outliers, it is better to have a larger sample size.
(2) The standard deviation of $\bar{d}$ is a measure of the **accuracy of the process** of using a sample mean difference to estimate the population mean difference.

$$s.d.(\bar{d}) = \frac{\sigma_d}{\sqrt{n}}$$

(3) In practice, the population standard deviation $\sigma_d$ is rarely known, so the sample standard deviation $s_d$ is used instead. When making this substitution we call the result a **standard error**.

---

**Standard error of the sample mean difference is given by:**

$$\textbf{s.e.}(\,\bar{d}\,) = \frac{s_d}{\sqrt{n}}$$

We can *interpret* **the standard error of the sample mean difference** as <u>estimating</u>, **approximately**, the **average distance** of the possible $\bar{d}$ value<u>s</u> (for repeated samples of the same size $n$) from the **population mean difference** $\mu_d$.

---

Moreover, we can use this standard error of the sample mean difference to produce a range of values that we are very confident will contain the population mean difference $\mu_d$, namely, $\bar{d} \pm$ (a few)s.e.($\bar{d}$). This is the basis for confidence interval for the population mean difference $\mu_d$, discussed in Part 2.

We will use the standard error of the sample mean difference to compute a standardized test statistic for testing hypotheses about the population mean difference $\mu_d$, namely,
<u>**Sample statistic – Null value.**</u>
**(Null) standard error**

This is the basis for testing about a population mean difference covered in Part 3.

**Additional Notes**

A place to … jot down questions you may have and ask during office hours, take a few extra notes, write out an extra problem or summary completed in lecture, create your own summary about these concepts.

# Stat 250 Gunderson Lecture Notes
## 8: Learning about a Population Mean Difference

### Part 2: Confidence Interval for a Population Mean of Paired Differences

## Confidence Interval for the Population Mean of Paired Differences $\mu_d$

Recall that an important special case of a single mean of a population occurs when two quantitative variables are **collected in pairs**, and we desire information about the **difference** between the two variables. For paired data designs, it is the *differences* that we are interested in analyzing. By focusing on the differences we again have just one sample of observations (the differences) and are able to use the confidence interval for the population mean difference.

**Notation:**

Population Parameter: $\mu_d$ = population mean difference

Data: $d_1, d_2, \ldots, d_n$

Sample Estimate: $\bar{d}$ = sample mean difference

Standard Error: s.e.($\bar{d}$) = $\frac{s_d}{\sqrt{n}}$ ($s_d$ = standard deviation of the sampled differences)

We use the sample estimate and its standard error to **form a confidence interval estimate** for the parameter using the following form:

### Sample Estimate ± Multiplier x Standard error

The multiplier used will depend on the confidence level, the sample size, and the type of parameter being estimated. In this case, since we are estimating a single population mean, the multiplier will be a $t*$ value. Here is the summary for a paired data confidence interval:

---

**One-sample $t$ Confidence Interval for the Population Mean Difference $\mu_d$**

$$\bar{d} \pm t^* \text{s.e.}(\bar{d})$$

where $t*$ is the appropriate value for a $t(n-1)$ distribution.

Note that $n$ is the number of pairs, or the number of differences. **This interval requires that the differences can be considered a random sample from a normal population.** If the sample size is large, the assumption of normality is not so crucial and the result is approximate.

---

## Try It!  Changes in Reasoning Scores

**Do piano lessons improve spatial-temporal reasoning of preschool children?**

**Data**: The change in reasoning score, after piano lessons - before piano lessons, with larger values indicating better reasoning, for a random sample of $n$ = 34 preschool children.

| 2 | 5 | 7 | -2 | 2 | 7 | 4 | 1 | 0 | 7 | 3 | 4 |
|---|---|---|----|---|---|---|---|---|---|---|----|
| 3 | 4 | 9 | 4 | 5 | 2 | 9 | 6 | 0 | 3 | 6 | -1 |
| 3 | 4 | 6 | 7 | -2 | 7 | -3 | 3 | 4 | 4 | | |

(a)  Display the data, summarize the distribution.
These data were entered into R to produce the following histogram.



Histogram of Change in Reasoning
(After-Before)
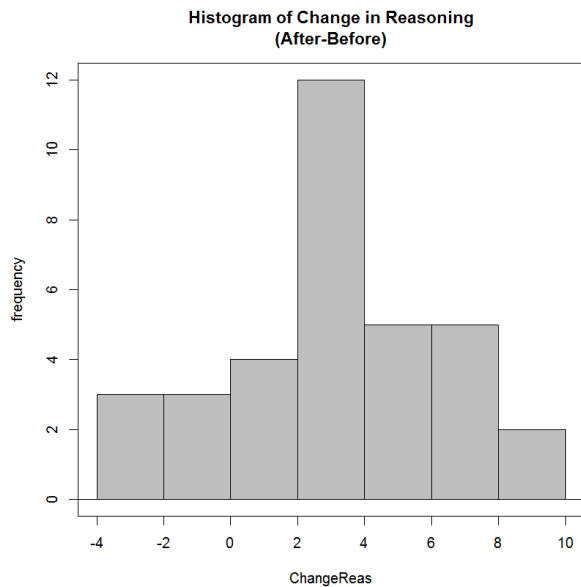
**Notes:**

1.  Diff = after – before so …

2. Sample mean difference = _____

3. Normality of the response
(the difference) for the population?

Some summary measures were obtained using R Commander and entered into the table:

```
> numSummary(Dataset[,"ChangeReas"], statistics=c("mean", "sd", "IQR",
+    "quantiles"), quantiles=c(0,.25,.5,.75,1))
    mean        sd IQR 0% 25% 50% 75% 100%  n
 3.617647 3.055196   4 -3   2   4   6    9 34
```

| Summary Statistics | | | |
|---|---|---|---|
| Mean diff ($\bar{d}$) | Std. Dev ($s_d$) | Sample size (n) | Std. Error |
| 3.62 | 3.06 | 34 | 0.52 |

(b)  Give a 95% confidence interval for the population mean improvement in reasoning scores.

(c)  What value is of particular interest to see whether or not it is in the interval?

(d) A student in your class wrote the following interpretation about the 95% confidence level used in making the interval. Is it a correct interpretation? If not, update it to make it correct.

*"If this study were repeated many times, we would expect 95% of the resulting confidence intervals to contain the sample mean improvement in reasoning scores."*

**R Note:**

The differences were already computed and entered as the data. So to make a confidence interval with R Commander we would need to perform a single-sample t-Test on the differences (and leave the null hypothesis value the default of 0). Be sure the confidence level is the one you want, namely .95.



R Single-Sample t-Test

Variable (pick one)
ChangeReas

Alternative Hypothesis
⦿ Population mean != mu0    Null hypothesis: mu =   0.0
◯ Population mean < mu0      Confidence Level:  .95
◯ Population mean > mu0

Help      Reset      OK      Cancel      Apply
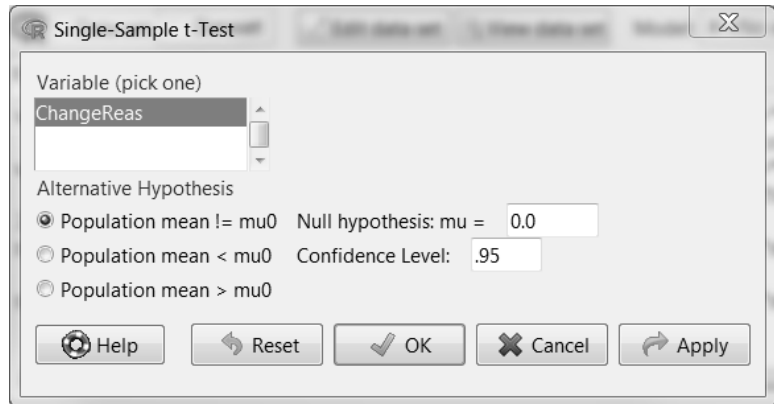
```
> with(Dataset, (t.test(ChangeReas, alternative='two.sided', mu=0.0,
+    conf.level=.95)))

        One Sample t-test

data:  ChangeReas
t = 6.9044, df = 33, p-value = 6.919e-08
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 2.551639 4.683655
sample estimates:
mean of x
 3.617647
```

| Paired T Results | | | |
|---|---|---|---|
| Mean diff ($\bar{d}$) | df | 95% CI Lower | 95% CI Upper |
| 3.62 | 33 | 2.55 | 4.68 |

If the before and after scores were entered into R, then we would use a paired t-test option. R would compute the differences for us and provide the confidence interval results. Details of the R steps for analyzing paired data can be found in your Lab Workbook.

## Additional Notes

A place to … jot down questions you may have and ask during office hours, take a few extra notes, write out an extra problem or summary completed in lecture, create your own summary about these concepts.

## Population Mean

| | |
|---|---|
| **Parameter** | $\mu$ |
| **Statistic** | $\bar{x}$ |

**Standard Error**

$$\text{s.e.}(\bar{x}) = \frac{s}{\sqrt{n}}$$

**Confidence Interval**

$$\bar{x} \pm t^* \text{s.e.}(\bar{x}) \qquad \qquad df = n - 1$$

**Paired Confidence Interval**

$$\bar{d} \pm t^* \text{s.e.}(\bar{d}) \qquad \qquad df = n - 1$$

**One-Sample $t$-Test**

$$t = \frac{\bar{x} - \mu_0}{\text{s.e.}(\bar{x})} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \qquad df = n - 1$$

**Paired $t$-Test**

$$t = \frac{\bar{d} - 0}{\text{s.e.}(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad df = n - 1$$

# Stat 250 Gunderson Lecture Notes
## 8: Learning about a Population Mean Difference

### Part 3: Testing about a Population Mean of Paired Difference

## Testing Hypotheses about the Population Mean of Paired Differences $\mu_d$

An important special case of a single mean of a population occurs when two quantitative variables are **collected in pairs**, and we desire information about the **difference** between the two variables. For paired data designs, it is the **_differences_** that we are interested in analyzing. By focusing on the differences we again have just one sample of observations (the differences) and are able to perform a one-sample t-test on the differences.

The procedure for the significance test is really the same - the sample mean of the data, $\bar{x}$ , is just written as $\bar{d}$ , primarily to emphasize that this is a paired design with the differences being analyzed. The commonly used null hypothesis is that the population mean difference is $\mu_d$ = 0.

**Possible null and alternative hypotheses.**

> 1. **H$_0$:**                              **versus H$_a$:**
>
> 2. **H$_0$:**                              **versus H$_a$:**
>
> 3. **H$_0$:**                              **versus H$_a$:**

Note:  The format of the alternative hypothesis depends on the research question of interest and the order in which the differences were taken.


**Test Statistic and Conditions for the Test**


Test statistic =  Sample statistic – Null value
                            Standard error

If H$_0$ is true, this test statistic has a _____ distribution.

We use this distribution to report the (bounds for the) *p*-value.


**Conditions for the test:**  The difference is assumed to be normally distributed for the population (but if the sample size is large, this condition is less crucial). So you need to examine the differences graphically and assess if there are any extreme outliers or skewness in the differences. If so, either the sample size needs to be large or an alternative testing method may be required.

## Try It! Knob Turning

A study involved *n*=25 right-handed students and a device with two different knobs (right-hand thread and left-hand thread). The **response** of interest is the time it takes to move knob indicator a fixed distance. The **question** of interest is to assess if right-hand threads are easier to turn on average. Use a 5% significance level.

a. Why is this a paired design and how should randomization be used in the experiment?


b. State the hypotheses. $H_0$: _____ versus $H_a$: _____

Here are a few summaries of each set of responses separately and then of the paired data:

**Paired Samples Statistics**

| | | Mean | N | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pair 1 | RTHREAD | 104.00 | 25 | 15.93 | 3.19 |
| | LTHREAD | 117.44 | 25 | 27.26 | 5.45 |

Below are the *t*-test results generated using **R Commander** and selecting **Statistics > Means > Paired T Test** and the correct direction for the alternative hypothesis. Notice that a 95% one-sided confidence bound is provided since our test alternative was one-sided to the left. If you wanted to also report a regular 95% confidence interval, you would run a two-sided hypothesis test in R.

| Summary Statistics | | | |
|---|---|---|---|
| **Mean diff ($\bar{d}$)** | **Std. Dev ($s_d$)** | **Sample size (n)** | **Std. Error** |
| -13.44 | 23.06 | 25 | 4.61 |

| Paired T Results | | | | |
|---|---|---|---|---|
| *t* | *df* | *p-value* | *95% CI Lower* | *95% CI Upper* |
| -2.914 | 24 | 0.004 | *** | -5.55 |

c. Perform the test.




d. Which are assumptions required for performing the paired t-test?
   • the turning times for the right-hand threaded knob are independent of the turning times for the left-hand threaded knob.
   • the turning times for the right-hand threaded knob are normally distributed.
   • the difference in turning times (diff = RT – LT) is normally distributed.

## Additional Notes

A place to … jot down questions you may have and ask during office hours, take a few extra notes, write out an extra problem or summary completed in lecture, create your own summary about these concepts.

| Population Mean | |
|---|---|
| **Parameter** | $\mu$ |
| **Statistic** | $\bar{x}$ |

**Standard Error**

$$\text{s.e.}(\bar{x}) = \frac{s}{\sqrt{n}}$$

**Confidence Interval**

$$\bar{x} \pm t^* \text{s.e.}(\bar{x}) \qquad \text{df} = n - 1$$

**Paired Confidence Interval**

$$\bar{d} \pm t^* \text{s.e.}(\bar{d}) \qquad \text{df} = n - 1$$

**One-Sample $t$-Test**

$$t = \frac{\bar{x} - \mu_0}{\text{s.e.}(\bar{x})} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \qquad \text{df} = n - 1$$

**Paired $t$-Test**

$$t = \frac{\bar{d} - 0}{\text{s.e.}(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad \text{df} = n - 1$$