Inference for Two Means

Nate Wells

Math 141, 4/21/21

Outline

In this lecture, we will. . .

- Investigate the theoretical distribution for difference in two means.
- Create confidence intervals and perform hypothesis tests using t distribution for differences in means.
- Compare inference procedures for two independent samples vs. paired samples

Section 1

Inference for 2 Means

Differences in Means

- Consider the following questions:

Differences in Means

- Consider the following questions:
  - Are variations in test scores between two sections of Math 141 just due to random sampling, or do they suggest an underlying difference?

Differences in Means

- Consider the following questions:
  - Are variations in test scores between two sections of Math 141 just due to random sampling, or do they suggest an underlying difference?
  - Does daily consumption of coffee improve cardiovascular health compared to a control?

Differences in Means

- Consider the following questions:
  - Are variations in test scores between two sections of Math 141 just due to random sampling, or do they suggest an underlying difference?
  - Does daily consumption of coffee improve cardiovascular health compared to a control?
  - Is there a significant difference in price between .99 and 1.0 carat diamonds?

Differences in Means

- Consider the following questions:
  - Are variations in test scores between two sections of Math 141 just due to random sampling, or do they suggest an underlying difference?
  - Does daily consumption of coffee improve cardiovascular health compared to a control?
  - Is there a significant difference in price between .99 and 1.0 carat diamonds?
- Each of these questions can be answered by analyzing the difference in means between samples taken from two groups.

Differences in Means

- Consider the following questions:
  - Are variations in test scores between two sections of Math 141 just due to random sampling, or do they suggest an underlying difference?
  - Does daily consumption of coffee improve cardiovascular health compared to a control?
  - Is there a significant difference in price between .99 and 1.0 carat diamonds?
- Each of these questions can be answered by analyzing the difference in means between samples taken from two groups.
- Groups could be formed from...

Differences in Means

- Consider the following questions:
    - Are variations in test scores between two sections of Math 141 just due to random sampling, or do they suggest an underlying difference?
    - Does daily consumption of coffee improve cardiovascular health compared to a control?
    - Is there a significant difference in price between .99 and 1.0 carat diamonds?
- Each of these questions can be answered by analyzing the difference in means between samples taken from two groups.
- Groups could be formed from. . .
    - Two different populations.
    - Two subsets within the same sample distinguished by levels of a categorical variable.
    - Two treatment groups in an experiment.

Distribution for Difference in Means

- Suppose random samples of size n_1 and n_2 are drawn **independentally** from populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.

Distribution for Difference in Means

- Suppose random samples of size n_1 and n_2 are drawn **independentally** from populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.
- **Goal**: Estimate the value of the parameter $\mu_1 - \mu_2$ using the statistic $\bar{x}_1 - \bar{x}_2$.

Distribution for Difference in Means

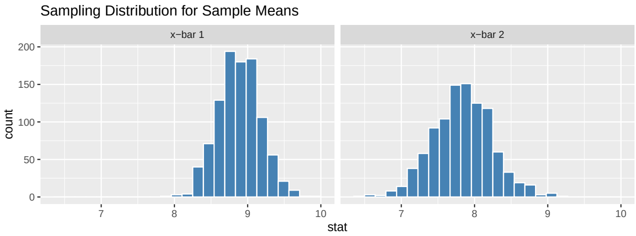
- Suppose random samples of size n_1 and n_2 are drawn **independentally** from populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.
- **Goal**: Estimate the value of the parameter $\mu_1 - \mu_2$ using the statistic $\bar{x}_1 - \bar{x}_2$.
    - We need to know the shape, center, and spread of distribution of $\bar{x}_1 - \bar{x}_2$.

Distribution for Difference in Means

- Suppose random samples of size n_1 and n_2 are drawn **independentally** from populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.
- **Goal**: Estimate the value of the parameter $\mu_1 - \mu_2$ using the statistic $\bar{x}_1 - \bar{x}_2$.
  - We need to know the shape, center, and spread of distribution of $\bar{x}_1 - \bar{x}_2$.
- By CLT, the distributions of $\bar{x}_1$ and $\bar{x}_2$ are approximately Normal.

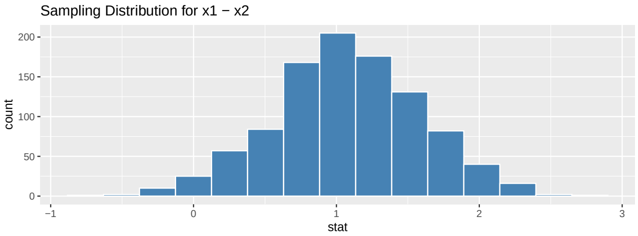
Distribution for Difference in Means

- Suppose random samples of size n_1 and n_2 are drawn **independentally** from populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.
- **Goal**: Estimate the value of the parameter $\mu_1 - \mu_2$ using the statistic $\bar{x}_1 - \bar{x}_2$.
  - We need to know the shape, center, and spread of distribution of $\bar{x}_1 - \bar{x}_2$.
- By CLT, the distributions of $\bar{x}_1$ and $\bar{x}_2$ are approximately Normal.



Distribution for Difference in Means

- Suppose random samples of size n_1 and n_2 are drawn **independentally** from populations with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.
- **Goal**: Estimate the value of the parameter $\mu_1 - \mu_2$ using the statistic $\bar{x}_1 - \bar{x}_2$.
  - We need to know the shape, center, and spread of distribution of $\bar{x}_1 - \bar{x}_2$.
- The distribtuion of the difference $\bar{x}_1 - \bar{x}_2$ is approximately Normal also



New Slide

- By the Central Limit Theorem, as both n_1 and n_2 get larger, the distribution of difference in sample means $\bar{x}_1 - \bar{x}_2$ becomes more Normally distributed, with mean $\mu_1 - \mu_2$ and standard error $SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

New Slide

- By the Central Limit Theorem, as both n_1 and n_2 get larger, the distribution of difference in sample means $\bar{x}_1 - \bar{x}_2$ becomes more Normally distributed, with mean $\mu_1 - \mu_2$ and standard error $SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
- In practice, we estimate the parameters σ_1 and σ^2 with the sampel statistics s_1 and s_2

New Slide

- By the Central Limit Theorem, as both n_1 and n_2 get larger, the distribution of difference in sample means $\bar{x}_1 - \bar{x}_2$ becomes more Normally distributed, with mean $\mu_1 - \mu_2$ and standard error $SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
- In practice, we estimate the parameters σ_1 and σ^2 with the sampel statistics s_1 and s_2
- Consider the standardized difference in sample means:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\mathrm{SE}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

New Slide

- By the Central Limit Theorem, as both n_1 and n_2 get larger, the distribution of difference in sample means $\bar{x}_1 - \bar{x}_2$ becomes more Normally distributed, with mean $\mu_1 - \mu_2$ and standard error $SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
- In practice, we estimate the parameters σ_1 and σ^2 with the sampel statistics s_1 and s_2
- Consider the standardized difference in sample means:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\mathrm{SE}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Theorem

The standardized difference t is approximately t-distributed with degrees of freedom $df = \min\{n_1 - 1, n_2 - 1\}$.

This approximation is appropriate either when both sample sizes are large (i.e. $n_1, n_2 \geq 30$*), or when both populations are approximately Normally distributed.*

Diamonds are Forever

Question: Do 1.0 carat diamonds command a higher price than .99 carat diamonds beyond what you would expect due to increase in weight?

- To answer, we collect random samples of 30 1.0 and 23 .99 carat diamonds.

Diamonds are Forever

Question: Do 1.0 carat diamonds command a higher price than .99 carat diamonds beyond what you would expect due to increase in weight?

- To answer, we collect random samples of 30 1.0 and 23 .99 carat diamonds.
  - To decouple the effect of increase in size between the two groups, we divide price of the .99 carat diamonds by .99 to obtain price per carat ppc

Diamonds are Forever

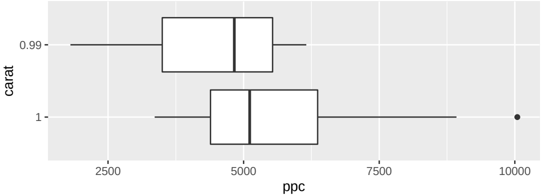
Question: Do 1.0 carat diamonds command a higher price than .99 carat diamonds beyond what you would expect due to increase in weight?

- To answer, we collect random samples of 30 1.0 and 23 .99 carat diamonds.
  - To decouple the effect of increase in size between the two groups, we divide price of the .99 carat diamonds by .99 to obtain price per carat ppc
- Here are side-by-side boxplots of the diamonds, along with summary statistics

Diamonds are Forever

Question: Do 1.0 carat diamonds command a higher price than .99 carat diamonds beyond what you would expect due to increase in weight?

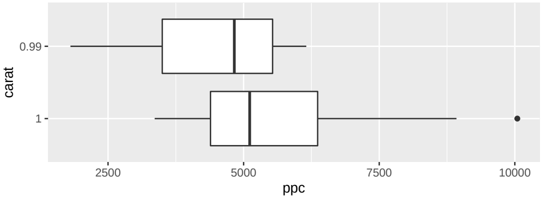
- To answer, we collect random samples of 30 1.0 and 23 .99 carat diamonds.
  - To decouple the effect of increase in size between the two groups, we divide price of the .99 carat diamonds by .99 to obtain price per carat ppc
- Here are side-by-side boxplots of the diamonds, along with summary statistics



Diamonds are Forever

Question: Do 1.0 carat diamonds command a higher price than .99 carat diamonds beyond what you would expect due to increase in weight?

- To answer, we collect random samples of 30 1.0 and 23 .99 carat diamonds.
  - To decouple the effect of increase in size between the two groups, we divide price of the .99 carat diamonds by .99 to obtain price per carat ppc
- Here are side-by-side boxplots of the diamonds, along with summary statistics



carat	mean_ppc	sd_ppc	n
0.99	4451	1332	23
1	5585	1614	30

Hypothesis Test

1. Our null and alternate hypotheses:

$$H_0 : \mu_1 - \mu_{99} = 0 \qquad H_0 : \mu_1 - \mu_{99} > 0$$

Hypothesis Test

1. Our null and alternate hypotheses:

$$H_0 : \mu_1 - \mu_{99} = 0 \qquad H_0 : \mu_1 - \mu_{99} > 0$$

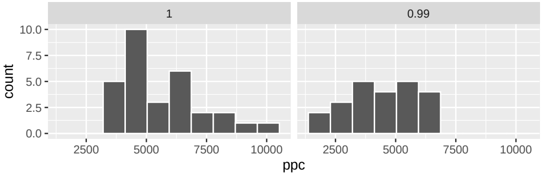
2. Assess conditions:

Hypothesis Test

1. Our null and alternate hypotheses:

$$H_0 : \mu_1 - \mu_{99} = 0 \qquad H_0 : \mu_1 - \mu_{99} > 0$$

2. Assess conditions:

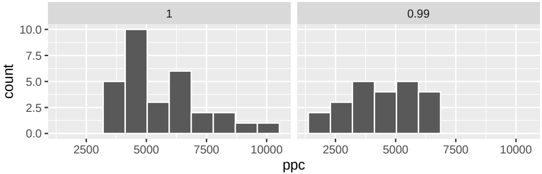


Hypothesis Test

1. Our null and alternate hypotheses:

$$H_0 : \mu_1 - \mu_{99} = 0 \qquad H_0 : \mu_1 - \mu_{99} > 0$$

2. Assess conditions:



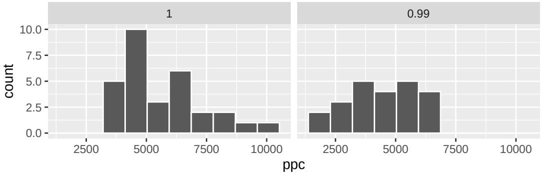
- The 1.0 carat diamonds show evidence of right-skew, but since $n \geq 30$, this is *probably* fine.

Hypothesis Test

1. Our null and alternate hypotheses:

$$H_0 : \mu_1 - \mu_{99} = 0 \qquad H_0 : \mu_1 - \mu_{99} > 0$$

2. Assess conditions:



- The 1.0 carat diamonds show evidence of right-skew, but since $n \geq 30$, this is *probably* fine.
- By construction, the two samples are independent. And observations within each sample are independent as well.

Hypothesis Test II

3. We compute our test statistic

Hypothesis Test II

3. We compute our test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{5585 - 4451}{\sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}} = 2.802$$

Hypothesis Test II

3. We compute our test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{5585 - 4451}{\sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}} = 2.802$$

4. Calculate the P-Value.

Hypothesis Test II

3. We compute our test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{5585 - 4451}{\sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}} = 2.802$$

4. Calculate the P-Value.

- The theory-based method says t is approximately t-distributed with 22 degrees of freedom: $\mathrm{df} = \min(23-1, 30-1) = 22$

Hypothesis Test II

3. We compute our test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{5585 - 4451}{\sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}} = 2.802$$

4. Calculate the P-Value.

- The theory-based method says t is approximately t-distributed with 22 degrees of freedom: $\mathrm{df} = \min(23-1, 30-1) = 22$

```
p_value<-1 - pt( 2.802, df = 22)
p_value
```

```
## [1] 0.0052
```

Hypothesis Test II

3. We compute our test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{5585 - 4451}{\sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}} = 2.802$$

4. Calculate the P-Value.

- The theory-based method says t is approximately t-distributed with 22 degrees of freedom: $\mathrm{df} = \min(23-1, 30-1) = 22$

```
p_value<-1 - pt( 2.802, df = 22)
p_value
```

```
## [1] 0.0052
```

5. Conclude

Hypothesis Test II

3. We compute our test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{5585 - 4451}{\sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}} = 2.802$$

4. Calculate the P-Value.

- The theory-based method says t is approximately t-distributed with 22 degrees of freedom: $\mathrm{df} = \min(23 - 1, 30 - 1) = 22$

```
p_value<-1 - pt( 2.802, df = 22)
p_value
```

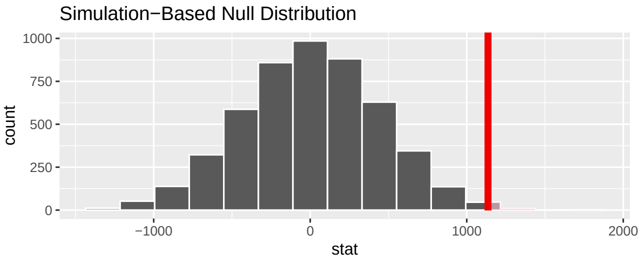
```
## [1] 0.0052
```

5. Conclude

- At the $\alpha = 0.01$ significance level, we reject the null hypothesis. This sample suggests 1.0 carat diamonds command a higher price than is explained by increase in weight alone.

Comparison with infer

```
set.seed(101)
diamonds_null <- diamonds %>% specify(ppc ~ carat) %>%
  hypothesize(null = "independence") %>%
  generate(reps = 5000, type = "permute") %>%
  calculate(stat = "diff in means", order = c("1", "0.99"))
diamonds_null %>% visualize()+shade_p_value(obs_stat = 1135, direction = "right")
```




```
diamonds_null %>% get_p_value(obs_stat = 1135, direction = "right" )
```

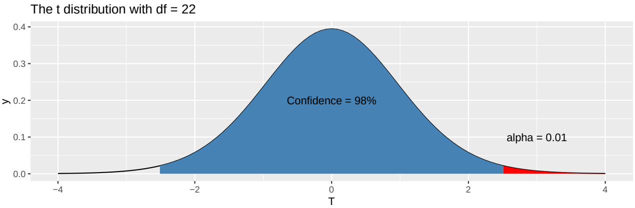
```
## # A tibble: 1 x 1
##   p_value
##     <dbl>
## 1  0.0042
```

An Equivalent Confidence Interval

Goal: Create a confidence interval that corresponds to **one-sided** $\alpha = 0.01$ significance

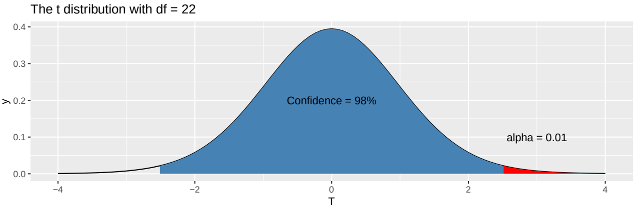
An Equivalent Confidence Interval

Goal: Create a confidence interval that corresponds to **one-sided** $\alpha = 0.01$ significance



An Equivalent Confidence Interval

Goal: Create a confidence interval that corresponds to **one-sided** $\alpha = 0.01$ significance



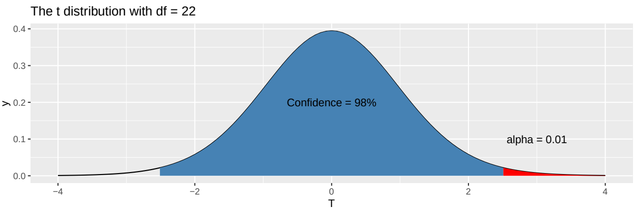
What is the t^* critical value for 98% confidence?

```
t_star<- qt(.99, df = 22)
t_star
```

```
## [1] 2.5
```

An Equivalent Confidence Interval

Goal: Create a confidence interval that corresponds to **one-sided** $\alpha = 0.01$ significance



What is the t^* critical value for 98% confidence?

```
t_star<- qt(.99, df = 22)
t_star
```

```
## [1] 2.5
```

Note that our observed t statistic was $t = 2.802$, which was more extreme than the critical value for 98% confidence

Confidence Interval

Create a 98% confidence interval to estimate the difference $\mu_1 - \mu_{99}$

Confidence Interval

Create a 98% confidence interval to estimate the difference $\mu_1 - \mu_{99}$

The formula for our confidence interval is

$$(\bar{x}_1 - \bar{x}_{99}) \pm t^* \cdot SE$$

Confidence Interval

Create a 98% confidence interval to estimate the difference $\mu_1 - \mu_{99}$

The formula for our confidence interval is

$$(\bar{x}_1 - \bar{x}_{99}) \pm t^* \cdot SE$$

Giving an interval of

$$(5585 - 4451) \pm 2.508 \cdot \sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}$$

or

$$(118.42, 2149.58)$$

Confidence Interval

Create a 98% confidence interval to estimate the difference $\mu_1 - \mu_{99}$

The formula for our confidence interval is

$$(\bar{x}_1 - \bar{x}_{99}) \pm t^* \cdot SE$$

Giving an interval of

$$(5585 - 4451) \pm 2.508 \cdot \sqrt{\frac{1614^2}{30} + \frac{1332^2}{23}}$$

or

$$(118.42, 2149.58)$$

Since this interval **does not** contain 0, we conclude that there IS a price increase for 1.0 carat diamonds.

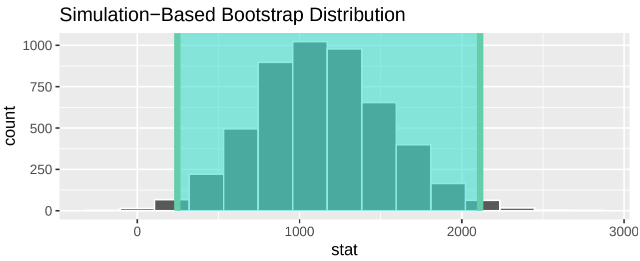
Comparison with infer

```
set.seed(101)
diamonds_boot <- diamonds %>% specify(ppc ~ carat) %>%
  generate(reps = 5000, type = "bootstrap") %>%
  calculate(stat = "diff in means", order = c("1", "0.99"))
```

```
diamonds_ci <- diamonds_boot %>% get_ci(level = 0.98, type = "percentile")
diamonds_ci
```

```
## # A tibble: 1 x 2
##   lower_ci upper_ci
##      <dbl>    <dbl>
## 1     247.    2113.
```

```
diamonds_boot %>%visualize()+ shade_ci(endpoints = diamonds_ci)
```



Section 2

Inference for Paired Samples

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.
  - It *is* possible that any observed effect is not due to the explanatory variable, but to some confounding variable present in one sample but not other.

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.
  - It *is* possible that any observed effect is not due to the explanatory variable, but to some confounding variable present in one sample but not other.
- To reduce effect of confounding variable, you could match individuals in one sample with similar individuals in the other sample.

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.
  - It *is* possible that any observed effect is not due to the explanatory variable, but to some confounding variable present in one sample but not other.
- To reduce effect of confounding variable, you could match individuals in one sample with similar individuals in the other sample.
  - But if matching is used in sample design, it is **not** appropriate to use the 2 sample procedures. (Why?)

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.
  - It *is* possible that any observed effect is not due to the explanatory variable, but to some confounding variable present in one sample but not other.
- To reduce effect of confounding variable, you could match individuals in one sample with similar individuals in the other sample.
  - But if matching is used in sample design, it is **not** appropriate to use the 2 sample procedures. (Why?)
- You can create a new variable recording the **difference** in measurements in each pair of individuals

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.
  - It *is* possible that any observed effect is not due to the explanatory variable, but to some confounding variable present in one sample but not other.
- To reduce effect of confounding variable, you could match individuals in one sample with similar individuals in the other sample.
  - But if matching is used in sample design, it is **not** appropriate to use the 2 sample procedures. (Why?)
- You can create a new variable recording the **difference** in measurements in each pair of individuals
- This new variable can be used to perform statistical inference using the 1-sample procedures for mean.

Matched Pairs

- Suppose you intend to design an experiment to determine whether the mean of two populations are equal.
- You could obtain an SRS from each population, compute means for each sample, take the difference, and assess variability based on previous procedures.
  - It *is* possible that any observed effect is not due to the explanatory variable, but to some confounding variable present in one sample but not other.
- To reduce effect of confounding variable, you could match individuals in one sample with similar individuals in the other sample.
  - But if matching is used in sample design, it is **not** appropriate to use the 2 sample procedures. (Why?)
- You can create a new variable recording the **difference** in measurements in each pair of individuals
- This new variable can be used to perform statistical inference using the 1-sample procedures for mean.
  - Rather than looking at the difference in means, we look at the mean of differences!

The World's Fastest Swimsuit

In the 2008 Olympics, controversy erupted over whether a new swimsuit design provided an unfair advantage to swimmers. Eventually, the International Swimming Organization banned the new suit. But can a certain suit really make a swimmer faster?

Data

A study analyzed max velocities for 12 pro swimmers with and without the suit:

swimmer	with_suit	without_suit	difference
1	1.6	1.5	0.08
2	1.5	1.4	0.10
3	1.4	1.4	0.07
4	1.4	1.3	0.08
5	1.2	1.1	0.10
6	1.8	1.6	0.11
7	1.6	1.6	0.05
8	1.6	1.5	0.05
9	1.6	1.5	0.06
10	1.5	1.4	0.08
11	1.5	1.4	0.05
12	1.5	1.4	0.10

Data

A study analyzed max velocities for 12 pro swimmers with and without the suit:

swimmer	with_suit	without_suit	difference
1	1.6	1.5	0.08
2	1.5	1.4	0.10
3	1.4	1.4	0.07
4	1.4	1.3	0.08
5	1.2	1.1	0.10
6	1.8	1.6	0.11
7	1.6	1.6	0.05
8	1.6	1.5	0.05
9	1.6	1.5	0.06
10	1.5	1.4	0.08
11	1.5	1.4	0.05
12	1.5	1.4	0.10

- Without performing any statistical inference, what is the likely conclusion to draw from this data?

The Setup

We want to determine whether the *average* max velocity with the suit is larger than the average without the suit.

The Setup

We want to determine whether the *average* max velocity with the suit is larger than the average without the suit.

- What is wrong with the following infer code? (If you try to run it, you'll get an error)

```
swim %>% specify(with_suit ~ without_suit) %>%
  hypothesize(null = "independence") %>%
  generate(reps = 1000, type = "permute") %>%
  calculate(stat = "diff in means", order = c("with_suit", "without_suit"))
```

The Setup

We want to determine whether the *average* max velocity with the suit is larger than the average without the suit.

- What is wrong with the following infer code? (If you try to run it, you'll get an error)

```
swim %>% specify(with_suit ~ without_suit) %>%
  hypothesize(null = "independence") %>%
  generate(reps = 1000, type = "permute") %>%
  calculate(stat = "diff in means", order = c("with_suit", "without_suit"))
```

- Review the data again...

swimmer	with_suit	without_suit	difference
1	1.6	1.5	0.08
2	1.5	1.4	0.10
3	1.4	1.4	0.07
4	1.4	1.3	0.08
5	1.2	1.1	0.10
6	1.8	1.6	0.11

The Setup, redux

We want to determine whether the *average* difference in max velocity (with - without) is positive. Let μ be the average difference.

$$H_0 : \mu = 0 \quad H_a : \mu > 0$$

The Setup, redux

We want to determine whether the *average* difference in max velocity (with - without) is positive. Let μ be the average difference.

$$H_0 : \mu = 0 \quad H_a : \mu > 0$$

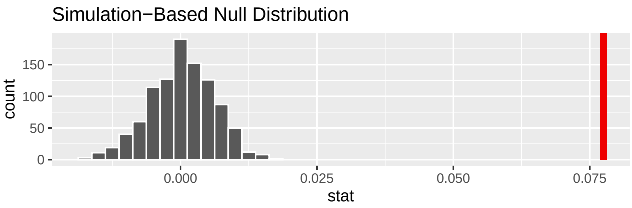
```
set.seed(1234)
swim_stat <- swim %>% specify(response = difference) %>%
  calculate(stat = "mean")
swim_stat
```

```
## # A tibble: 1 x 1
##      stat
##     <dbl>
## 1 0.0775
```

```
swim_null <- swim %>% specify(response = difference) %>%
  hypothesize(null = "point", mu = 0) %>%
  generate(reps = 1000, type = "bootstrap") %>%
  calculate(stat = "mean")
```

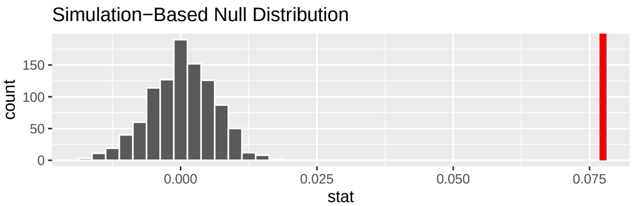
Conclusion

```
swim_null %>% visualise()+shade_p_value(obs_stat = swim_stat, direction = "right")
```



Conclusion

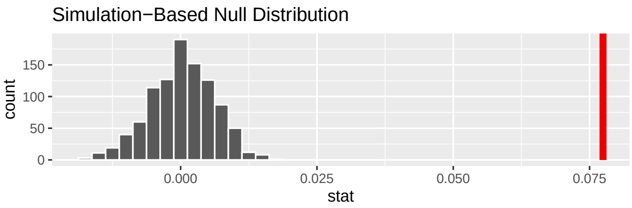
```
swim_null %>% visualise()+shade_p_value(obs_stat = swim_stat, direction = "right")
```



- P-Value?

Conclusion

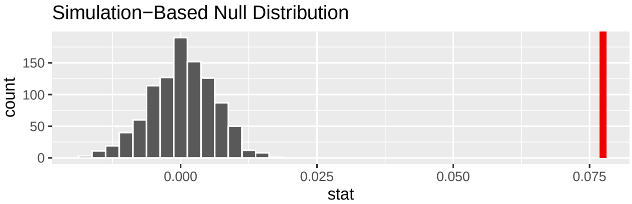
```
swim_null %>% visualise()+shade_p_value(obs_stat = swim_stat, direction = "right")
```



- P-Value?
- This study gives good evidence at the $\alpha = 0.01$ significance level that the suit does increase max velocity.

Conclusion

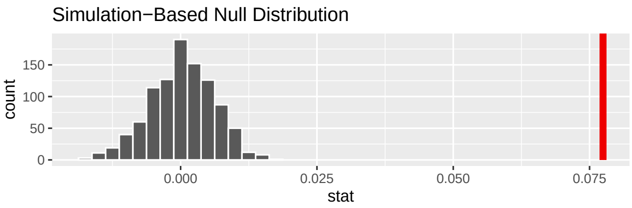
```
swim_null %>% visualise()+shade_p_value(obs_stat = swim_stat, direction = "right")
```



- P-Value?
- This study gives good evidence at the $\alpha = 0.01$ significance level that the suit does increase max velocity.
- Is a difference of 0.077 max velocity of practical significance?

Conclusion

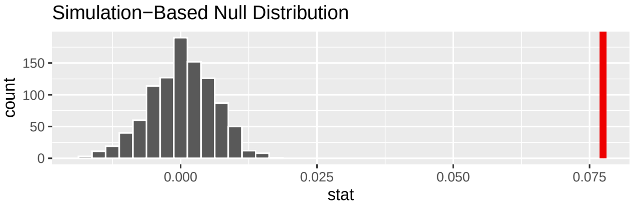
```
swim_null %>% visualise()+shade_p_value(obs_stat = swim_stat, direction = "right")
```



- P-Value?
- This study gives good evidence at the $\alpha = 0.01$ significance level that the suit does increase max velocity.
- Is a difference of 0.077 max velocity of practical significance?
  - If the average max velocity is about 1.4, this is about a 5% increase in speed.

Conclusion

```
swim_null %>% visualise()+shade_p_value(obs_stat = swim_stat, direction = "right")
```



- P-Value?
- This study gives good evidence at the $\alpha = 0.01$ significance level that the suit does increase max velocity.
- Is a difference of 0.077 max velocity of practical significance?
  - If the average max velocity is about 1.4, this is about a 5% increase in speed.
- In the 2008 Beijing Olympics, swimmers wearing the suit. . .
  - Were awarded 98% of all medals (including 33 of 36 gold medals).
  - Represented 23 of the total 25 world records broken.