Student’s t Test

- When the population standard deviation is not known, you cannot use a z score hypothesis test
- Use Student’s t test instead
- Student’s t, or t test is, conceptually, very similar to the z-score test we have been using, but uses an estimate of σ based on s
The t Test

When the standard deviation of the population is not known, as is usually the case, we must estimate the standard deviation of the population.

We use the standard deviation of the sample to estimate the population standard deviation: \( \sigma \approx s \)

Estimated Standard Error

\[
\sigma_M = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}}
\]

\[
s_M = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}}
\]

The t statistic is used to test hypotheses about an unknown population mean \( \mu \) when the value of \( \sigma \) is unknown:

\[
t = \frac{M - \mu}{s_m} = \frac{M - \mu}{\sqrt{s^2/n}}
\]
Sample and Population Standard Deviations

- Even the unbiased estimate of the population standard deviation will be inexact when the sample size is small (< 30).
- The smaller the sample size is, the less precise the unbiased estimate of the population standard deviation will be.
- Because of this imprecision, it is inappropriate to use the normal distribution.

The t Distributions

- William Gosset created a series of distributions known as the t distributions.
- The t distributions approximate the normal distributions, but account for the imprecision in the estimation of $\sigma$.
- As a point estimate of $\sigma$, tends to be more precise the larger the sample size.
- Because the imprecision in the estimate depends on sample size, there are multiple t distributions, depending on the degrees of freedom.
Degrees of Freedom

- Degrees of freedom correspond to the number of scores that are free to take on any value after restrictions (such as the value of the mean) are placed on the set of data.
- E.g., if the mean of 5 data points is 0, then how many data points can take on any value and still have the mean equal 0?

| 1 | X₁ |
| 2 | X₂ |
| 3 | X₃ |
| 4 | X₄ |
| 5 | (X₁ + X₂ + X₃ + X₄) |
| Mean | 0 |

Degrees of Freedom

- 4 of the 5 numbers can take on any value
- But the fifth number must equal -1 times the sum of the other four for the mean to equal 0
- Thus n - 1 scores are free to vary
- In this case df = n - 1

Example: Step 1

- On average, do people with schizophrenia smoke more cigarettes than the population (μ = 3 per day)?
- A sample of 25 people with schizophrenia will be used
  - \( \bar{X} = 6, s^2 = 36 \)
- Step 1: Write the hypotheses and select an α level:
  - \( H_0: \mu \leq 3 \)
  - \( H_1: \mu > 3 \)
  - \( \alpha = .05 \)
Step 2

- Locate the critical region
  - You need to know the degrees of freedom
  - For a one-sample t-test, the degrees of freedom equal: \( df = n - 1 \)
    \( df = 25 - 1 = 24 \)
  - Determine whether the hypothesis test is one or two tailed
    - The hypothesis asks if people with schizophrenia smoke more cigarettes than average; thus we have a 1 tailed test

- From a table of critical t values, determine the critical region
  - \( \alpha = .05, df = 24, \) one tailed
  - \( t_{critical} = 1.711 \)
    - Any calculated t value greater than or equal to 1.711 will cause us to reject \( H_0 \)

Determine the Critical Value

- To determine the critical t value, consult a table of critical t values
- Find the column that is labeled with your \( \alpha \) level
  - Make sure you select the right number of tails (1 vs 2)
- Find the row that is labeled with your degrees of freedom
- The critical t value is at the intersection
Determine the Critical Value

If the table does not contain the desired degrees of freedom, use the critical t value for the *next smallest* degrees of freedom.

Step 3: Calculate the Test Statistic

Because we do not know the population variance ($\sigma^2$), we cannot calculate the standard error (standard deviation of the distribution of sample means).

We must estimate the standard error from the sample variance ($s^2$):

$$S_M = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{36}{25}} = 1.2$$

Step 3: Calculate the Test Statistic

Plug and chug the t test value:

$$t_{obs} = \frac{M - \mu}{S_M} = \frac{6 - 3}{1.2} = 2.50$$
Step 4: Make a Decision

- If the observed $t$ (the value you calculated) is in the tail(s) cut off by the critical $t$, then you can reject $H_0$.
- Because our observed $t$ (2.50) is in the tail cut off by the critical $t$ (1.711), we reject $H_0$ that people with schizophrenia smoke less than or equal to the population

Assumptions of $t$

- Values in sample must be independent observations
- Population must be normally distributed

Effect Size for $t$

- Cohen’s $d = \frac{\text{mean difference}}{\text{standard deviation}} = \frac{\mu_{\text{treatment}} - \mu_{\text{control}}}{\sigma}$
- Estimated $d = \frac{\text{mean difference}}{\text{standard deviation}} = \frac{M - \mu}{s}$

<table>
<thead>
<tr>
<th>Magnitude of $d$</th>
<th>Evaluation of Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.2 \leq d &lt; 0.5$</td>
<td>Small effect</td>
</tr>
<tr>
<td>$0.5 \leq d &lt; 0.8$</td>
<td>Medium effect</td>
</tr>
<tr>
<td>$0.8 \leq d$</td>
<td>Large effect</td>
</tr>
</tbody>
</table>
$r^2$

$r^2$ is an alternative to Cohen’s $d$; it measures the proportion of variation in the scores that is explained by the treatment.

---

Example

- Nine infants see pictures of two faces – one rated as more attractive than the other.
- Infants look for 20 seconds – measure how long they look at the more attractive face.
  - $\bar{X}=13$ seconds, $SS (\Sigma(X-\mu)^2)=72$
  - $H_0$: $\mu_{\text{attractive}} = 10$

---

$r^2$

Adjusted scores, treatment effect removed

Original scores, including treatment effect

No effect $\mu = 10$

No effect $\mu = 10$
### Calculation of SS

<table>
<thead>
<tr>
<th>Score</th>
<th>Deviation from $\mu = 10$</th>
<th>Squared Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>17</td>
<td>7</td>
<td>49</td>
</tr>
</tbody>
</table>

SS = 153

### Calculation of SS after the treatment effect is removed

<table>
<thead>
<tr>
<th>Score</th>
<th>Adjusted Score</th>
<th>Deviation from $\mu = 10$</th>
<th>Squared Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.3</td>
<td>-5</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>10.3</td>
<td>-3</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>12.3</td>
<td>-1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>13.3</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15.3</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>17.3</td>
<td>4</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

SS = 72

### $r^2$

- How much variability was explained by the treatment?
- Total variability = Variability left after the effect of the treatment has been removed
- $SS_{\text{total}} - SS_{\text{after treatment removed}} = 153 - 72 = 81$
- $r^2 = \text{variability explained by treatment} / \text{total variability}$
- $r^2 = SS_{\text{treatment}} / SS_{\text{total}} = 81 / 153 = 0.5294$

### In this example

- $t = 3$ and $df = 8$ (see pages 292 – 294 of the textbook)
- $r^2 = \frac{3^2}{(3^2 + 8)} = \frac{9}{9 + 8} = \frac{9}{17} = 0.5294$

### Percentage of Variance Explained, $r^2$

<table>
<thead>
<tr>
<th>$r^2$</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01 ≤ $r^2 &lt; .09$</td>
<td>Small effect</td>
</tr>
<tr>
<td>.09 ≤ $r^2 &lt; .25$</td>
<td>Medium effect</td>
</tr>
<tr>
<td>.25 ≤ $r^2$</td>
<td>Large effect</td>
</tr>
</tbody>
</table>
Influence of n and $s^2$

- The value of $t$ is influenced by the sample size ($n$) and sample variance ($s^2$)
  - As $n$ increases, $s_M$ decreases and $t$ increases
  - As $s^2$ decreases, $s_M$ decreases and $t$ increases

$$t = \frac{M - \mu}{s_M} \quad s_M = \sqrt{\frac{s^2}{n}}$$

Confidence Interval

- A confidence interval is an interval estimate (e.g. a range of values used as an estimate of an unknown quantity) that is accompanied by a specific level of confidence (or probability)
  - E.g. There is a 95% chance that the population mean is between 45 and 55

Confidence Interval

- Psychologists often report the 95% confidence interval (or 95% CI)
  - If a given study was performed 100 times, and a 95% confidence interval was calculate for each of those replications of the study, then 95% of those confidence intervals would contain the population parameter and 5% would not
When To Use Estimation

- Used after a hypothesis test that rejects $H_0$
- If you already know that an effect exists and want to find out how big the effect is
- If you want basic information about an unknown population

Estimation with the t Statistic

- Single-sample:
  - $\mu = M \pm t \cdot s_m$
  - The mean ($M$) and standard error ($s_m$) all come from the data

The value of $t$ comes from a table of critical $t$ values and depends on the degrees of freedom and confidence level (e.g., 95%).

For a $t$ distribution with $df = 15$, 95% of the area under the distribution is between $\pm 2.131$. 
Problem

- Do female, indoor cats weigh a different amount than female cats in general?
- $\bar{x}_{\text{indoor}} = 13$ pounds
- $s_{\text{indoor}} = 2$ pounds; $s^2 = 4$ pounds
- $n_{\text{indoor}} = 16$
- $\mu_{\text{cats}} = 10$ pounds

Step 1

- Step 1: Write the hypotheses and select an $\alpha$ level:
  - $H_0$: $\mu = 10$
  - $H_1$: $\mu \neq 10$
  - $\alpha = .05$

Step 2

- Locate the critical region
  - You need to know the degrees of freedom
  - For a one-sample t-test, the degrees of freedom equal: $df = n - 1$
    - $df = 16 - 1 = 15$
  - Determine whether the hypothesis test is one or two tailed
    - The hypothesis asks if the indoor cats weigh a different amount than cats in general; thus we have a two tailed test
Step 2

- From a table of critical t values, determine the critical region
  - \( \alpha = .05 \), df = 15, two tailed
  - \( t_{\text{critical}} = 2.131 \)
  - Any calculated t value greater than or equal to 2.131 or less than -2.131 will cause us to reject \( H_0 \)

Step 3: Calculate the Test Statistic

- Because we do not know the population variance \( (\sigma^2) \), we cannot calculate the standard error (standard deviation of the distribution of sample means)
- We must estimate the standard error from the sample variance \( (s^2) \)

\[
S_M = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{4}{16}} = 0.50
\]

Step 3: Calculate the Test Statistic

- Plug and chug the t test value

\[
t_{\text{obs}} = \frac{M - \mu}{S_M} = \frac{13 - 10}{0.5} = 6
\]
Step 4: Make a Decision

- If the observed $t$ (the value you calculated) is in the tails cut off by the critical $t$, then you can reject $H_0$.
- Because our observed $t$ (6) is in the tails cut off by the critical $t$ (2.131), we reject $H_0$ that the weight of female, indoor cats equals the weight of female cats in general.

Cohen’s $d$

\[ \text{estimated } d = \frac{\text{mean difference}}{\text{standard deviation}} = \frac{M - \mu}{s} \]

\[ \text{estimated } d = \frac{13 - 10}{2} = \frac{3}{2} = 1.5 \]

$r^2$

\[ r^2 = \frac{t^2}{t^2 + df} \]

\[ r^2 = \frac{6^2}{6^2 + 15} = \frac{36}{36 + 15} = .706 \]
95% Confidence Interval

\[ \mu = M \pm t \cdot s_M \]

\[ M = 13 + 2.131 \cdot 0.50 = 14.0655 \]

\[ M = 13 - 2.131 \cdot 0.50 = 11.9345 \]

95% CI: 11.9 to 14.1