Chapter 9: Introduction to the $t$ Statistic

First of all, you need to know who developed the $t$ statistic. His name was William S. Gossett, but he published under the pseudonym Student. (His employer wouldn’t let him publish under his own name.) Thus, the statistic is often called Student’s $t$. But that’s not the good part. Gossett actually had the world’s best job…even better than L’il Abner’s job as a mattress tester. Can you guess how he earned his living?

You’ve learned how to test hypotheses about means using $z$-scores, but you should also recognize that the test is fairly artificial, because it requires that you know $\sigma$ in order to compute $\sigma_X$:

$$z = \frac{\bar{X} - \mu}{\sigma_X}$$

Even though you would rarely know $\sigma$, you could estimate it by using the sample standard deviation. When you must use $s$ to estimate $\sigma$, however, you can no longer compute $\sigma_X$. Instead, you will now compute an estimated standard error:

$$s_X = \hat{\sigma}_X = \sqrt{\frac{s^2}{n}} = \frac{s}{\sqrt{n}}$$

When you create a standard score using the estimated standard error, you will have created a $t$ statistic:

$$t = \frac{\bar{X} - \mu}{s_X}$$

The $t$ statistic is used to test hypotheses about an unknown population mean $\mu$ in situations where the value of $\sigma$ is unknown. The formula for the $t$ statistic has the usual standard score structure, and thus closely resembles the $z$-score, but with the estimated standard error in the denominator.

Unlike the unit normal distribution, there is not a single $t$ distribution. Instead there is a family of $t$ distributions, determined by the degrees of freedom found in the denominator of the sample variance ($n - 1$). The $t$ distributions are unimodal and symmetrical, with a mean of 0. However, the tails are typically higher than those found in a normal distribution, except for very large sample sizes (approaching infinity), in which case the $t$ distribution is normal. As a result, the values of $t$ that cut off the upper and lower 2.5% of the distribution (using the typical $\alpha$-level and a two-tailed test) will vary. In order to determine those values, a new table is needed. I’ve highlighted the column that you will use exclusively ($\alpha = .05$, two-tailed).
Note that the $t$-distribution approaches the normal distribution with large sample size, because the critical value of $t$ for an infinite sample size is 1.96, which is the $z$-score that would cut off 2.5% of the normal distribution.

**Hypothesis Testing with the $t$ Statistic**

G&W provide an example using an experiment based on Scaife’s (1976) study of aversion to eye-spot patterns among insectivorous birds. Birds are placed in a cage, which is divided in half. One side has eye spots and the other side does not. If the birds don’t try to avoid the eye spots, then they will stay in each half of the cage for an equal duration (the dv). Birds are placed in the cage for an hour, so they should spend about 30 minutes on each side of the cage.
\[ \text{H}_0: \mu_{\text{Plain Side}} = 30 \]
\[ \text{H}_1: \mu_{\text{Plain Side}} \neq 30 \]
\[ \alpha = 0.05 \]
\[ n = 16, \text{ so } t_{\text{Critical}}(15) = 2.131 \]

Decision rule: If \( |t_{\text{Obtained}}| \geq 2.131 \), reject \( \text{H}_0 \).

For this sample of 16 birds, the mean time spent on the plain side (\( \bar{X} \)) was 39 minutes. With \( SS = 540 \), you can compute the necessary values as follows:

\[
\begin{align*}
SS &= 540 \\
\sum s^2 &= \frac{540}{15} = 36 \\
\frac{s^2}{n} &= \frac{36}{16} = 1.5
\end{align*}
\]

You can now compute the \( t \) statistic as follows:

\[
t = \frac{39 - 30}{1.5} = 6
\]

You would reject \( \text{H}_0 \), because \( |t_{\text{Obtained}}| \geq t_{\text{Critical}} \)

You could then conclude that the birds spend significantly more time on the plain side of the cage than on the side with the eye spots.

You might report your results as follows:
The birds spent more time on the plain side of the apparatus (\( M = 39, SD = 6 \)), \( t(15) = 6.0, p < .05 \).

Effect Size

Effect size could be measured with Cohen’s \( d \) as follows:

\[
d = \frac{\text{mean difference}}{\text{sample standard deviation}} = \frac{9}{6} = 1.5
\]

This effect size would be quite large. (As a rough guide, \( d = 0.2 \) or less is a small effect size, a \( d \) between \( 0.2 \) and \( 0.8 \) is a medium effect size, and a \( d \) greater than \( 0.8 \) is a large effect size.)

Another measure of effect size is \( r^2 \):

\[
r^2 = \frac{t^2}{t^2 + df} = \frac{6^2}{6^2 + 15} = 0.71
\]

For \( r^2 \), a small effect is between \( 0.01 \) and \( 0.09 \), a medium effect is between \( 0.09 \) and \( 0.25 \), and a large effect is \( 0.25 \) or greater. Thus, using either \( d \) or \( r^2 \), this effect is quite large.
QUESTION: Do women’s salaries differ from men’s? Assume that we know that the mean salary for men is $41k (2003 census, $40,668).

STATISTICAL HYPOTHESES:  

\[ H_0: \mu = 41k \]
\[ H_1: \mu \neq 41k \]

DECISION RULE: Reject \( H_0 \) if \( t_{\text{obs}} \) occurs with \( p < .05 \).

With a sample \( n = 9, \, df = 8, \, t_{\text{crit}} = 2.306 \).

PICTURE:

DATA:

28
20
35
37
39
30
31
29
28

FORMULAS:

\[ t_{\text{obs}} = \frac{\overline{X} - \mu}{S_X} \]
\[ S_X = \frac{s}{\sqrt{n}} \]
\[ d = \frac{\overline{X} - \mu}{s} \]

DECISION:

CONCLUSION/INTERPRETATION:
A sample of 25 reaction time scores was selected from a group of professional athletes to see if their reflexes were faster than those found in the normal population. The sample had a mean of 10 sec and a sum of squares of 600. On the basis of years of study of normal, randomly selected people, we know that the distribution is normal with \( \mu \) of 13 sec. How likely is it that the sample was obtained from the normal population?

How would your computations differ if you were told that the sample size was 100?

If you were told that the standard deviation of the population (\( \sigma \)) was 2, what would you do differently from part 1?

When individuals feel disconnected from others in society, we say that they feel alienated. Dr. Conrad, a clinical psychologist, measures alienation with a survey scale that consists of 20 items, where the higher the score, the greater the feeling of alienation. The scale is normalized so that a score of 40 or below is considered healthy and not alienated. For 25 patients who are seeking care for anxiety, he finds a mean alienation score of 46 with a standard deviation of 13.2. Test the hypothesis that the sample was drawn from a population of people with a healthy alienation score of 40. [Ritchey, 332]

You have a measurement index of happiness with scores ranging from 1 to 10. You wish to test the hypothesis that the mean score of the Elmsville area is 6.5. You draw a random sample of 59 individuals and find a mean score of 5.9 with a standard deviation of 2.6. Test the hypothesis. [Ritchey, 332]
On a standardized spatial skills task, normative data reveal that people typically get $\mu = 15$ correct solutions. A psychologist tests $n = 7$ individuals who have brain injuries in the right cerebral hemisphere. For the following data, determine whether or not right-hemisphere damage results in significantly different performance on the spatial skills task.

<table>
<thead>
<tr>
<th>X</th>
<th>$X^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>144</td>
</tr>
<tr>
<td>16</td>
<td>256</td>
</tr>
<tr>
<td>9</td>
<td>81</td>
</tr>
<tr>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>289</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Sum</td>
<td>82</td>
</tr>
</tbody>
</table>

A psychologist is interested in the long-term effects of a divorce on the children in a family. A sample is obtained of $n = 10$ children whose parents were divorced at least five years ago. Each child is given a personality questionnaire measuring depression. In the general population, the scores on this questionnaire form a normal distribution with $\mu = 80$ and $\sigma = 12$. The scores for this sample are seen below. Do children from divorced families appear to differ in levels of depression from children in the general population?

<table>
<thead>
<tr>
<th>X</th>
<th>$X^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>6889</td>
</tr>
<tr>
<td>81</td>
<td>6561</td>
</tr>
<tr>
<td>75</td>
<td>5625</td>
</tr>
<tr>
<td>92</td>
<td>8464</td>
</tr>
<tr>
<td>84</td>
<td>7056</td>
</tr>
<tr>
<td>107</td>
<td>11449</td>
</tr>
<tr>
<td>63</td>
<td>3969</td>
</tr>
<tr>
<td>112</td>
<td>12544</td>
</tr>
<tr>
<td>92</td>
<td>8464</td>
</tr>
<tr>
<td>88</td>
<td>7744</td>
</tr>
<tr>
<td>Sum</td>
<td>877</td>
</tr>
</tbody>
</table>
Using SPSS to Conduct Single-Sample t-Tests

Single-Sample t-Test problem from Gravetter & Wallnau, 9th Edition (Demo 9.1)

First set up the data (in Variable View), as seen below:

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Width</th>
<th>Decimals</th>
<th>Label</th>
<th>Values</th>
<th>Missing</th>
<th>Columns</th>
<th>Align</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>optimism</td>
<td>8</td>
<td>2</td>
<td>None</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Next, enter the data (in Data View) as seen below left:

Once the data are entered, click-and-drag the Analyze menu to Compare Means, and then One-Sample T Test. You’ll then see a window like the one above on the right. Click-and-drag the optimism variable from the left into the box on the right. Note that to test H₀: µ = 15, I’ve entered 15 into the Test Value window. Click on OK when you have done so and you will see an analysis window opened. The analysis should appear, as seen below. Note that there is no need to use the table of critical values of t, because the program shows the exact probability of obtaining a t of -4.376 (in this case .002). All you have to do is determine if the probability level is greater than .05 (retain H₀) or less than .05 (reject H₀). In this case, you would reject H₀: µ = 15 and conclude that the optimism scores were significantly lower in this class.

<table>
<thead>
<tr>
<th>One-Sample Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>optimism</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Std. Deviation</td>
</tr>
<tr>
<td>Std. Error Mean</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>One-Sample Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Value = 15</td>
</tr>
<tr>
<td>t</td>
</tr>
<tr>
<td>df</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td>Mean Difference</td>
</tr>
<tr>
<td>95% Confidence Interval of the Difference</td>
</tr>
<tr>
<td>Lower</td>
</tr>
<tr>
<td>Upper</td>
</tr>
</tbody>
</table>
A sample of quiz scores was obtained from a statistics class. The SPSS analysis conducted on these data is seen below.

### One-Sample Statistics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz Scores</td>
<td>36</td>
<td>7.4444</td>
<td>1.36161</td>
<td>0.2593</td>
</tr>
</tbody>
</table>

### One-Sample Test

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz Scores</td>
<td>-2.448</td>
<td>35</td>
<td>0.020</td>
<td>-5.556</td>
<td>-10.163 to -0.949</td>
</tr>
</tbody>
</table>

a. State the null and alternative hypothesis.

b. How many quiz scores are involved in this analysis?

c. What is your decision about \( H_0 \)? What should you conclude?

d. How would you describe the effect size of this analysis?