Section 7.2 Homework Answers

7.53 A multimedia program designed to improve dietary behavior among low-income women was evaluated by comparing women who were randomly assigned to intervention and control groups. The intervention was a 30-minute session in a computer kiosk in the Food Stamp office.\(^2\) One of the outcomes was the score on a knowledge test taken about 2 months after the program.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>(\bar{x})</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>165</td>
<td>5.08</td>
<td>1.15</td>
</tr>
<tr>
<td>Control</td>
<td>212</td>
<td>4.33</td>
<td>1.16</td>
</tr>
</tbody>
</table>

(a) The test had six multiple-choice items that were scored as correct or incorrect, so the total score was an integer between 0 and 6. Do you think that these data are normally distributed? Explain why or why not.

No. First of all, a normally distributed population is continuous, which means the actual measurements we take for each observation can be any real number within some reasonable range of values. For example, human height can be measured in inches, feet, centimeter, or any other measure of distance. Except for the fact that our instruments only have so much accuracy, we recognize that we tend to view human height as being a continuous measure. If we had a measuring instrument that could measure to the billionth of an inch, then we could measure down to this accuracy; if we had an even more sensitive instrument we could get even more accuracy. Thus any real number is possible within some reasonable specified range (2 feet to 8.5 feet?).

Our measure consists of exactly 7 values: \(\{0, 1, 2, 3, 4, 5, 6\}\); thus, this is clearly not continuous. As a matter of fact we could rework our measure so we have a binomial situation which leads to proportions (after all the test is how many you get correct out of six questions).
(b) Is it appropriate to use the two-sample t procedures that we studied in this section to analyze these data? Give reasons for your answer.

The two-sample t-test is a robust test for large enough sample sizes. It will actually produce good results (i.e. the quoted probabilities, p-value, type I error, power, type II error will have the stated long run averages). The two samples sizes are 165 and 212, which is considered large enough.

Secondly, there is really no extreme values since we are limited to 7 possible values that are close to each other.

(c) Describe appropriate null and alternative hypotheses for evaluating the intervention. Some people would prefer a two-sided alternative in this situation while others would use a one-sided significance test. Give reasons for each point of view.

The researchers are trying to improve dietary behavior. While not stated, I am imagining that the 30-minute session with the computer is attempting to improve this behavior. Thus, the test should be a one-sided test.

Otherwise, we would do a two-sided test. In a two-sided test you are suggesting that either method, 30-minute session or no intervention at all could potentially be better. In this case, you are testing to see if the witnessed difference is large enough to suggest there is some difference.

(d) Carry out the significance test. Report the test statistic with the degrees of freedom and the P-value. Write a short summary of your conclusion.

\[ H_0: \mu_{\text{control}} = \mu_{\text{comp}} \quad H_0: \mu_{\text{control}} < \mu_{\text{comp}}. \]

\[ t = \frac{5.08 - 4.33}{\sqrt{\frac{1.15^2}{165} + \frac{1.16^2}{212}}} \]

\[ t = 6.258 \text{ using the “lesser of” rule my degrees of freedom is 164.} \]

\[ P(t > 6.258) = 1.63 \times 10^{-9} \text{ clearly a significant result. A difference as large as 5.08 – 4.33 = 0.75 is very unusual assuming no difference with such a large sample size.} \]

I used Excel to find the p-value above: I typed “tdist(6.258,164,1)”
(e) Find a 95% confidence interval for the difference between the two means. Compare the information given by the interval with the information given by the significance test. 

Now that we know there is some difference, the next natural question is what is that difference? We use a confidence interval to estimate that difference.

\[
5.08 - 4.33 \pm 1.975 \sqrt{\frac{1.15^2}{165} + \frac{1.16^2}{212}}
\]

To find the \( t^* \) multiplier I typed “tinv(0.05, 164)"

(f) The women in this study were all residents of Durham, North Carolina. To what extent do you think the results can be generalized to other populations?

Without much other information it would be difficult to extend beyond the residents of North Carolina. For example the demographics and economic landscape of Durham, North Carolina may be unique to that county and not applicable elsewhere.
Exposure to dust at work can lead to lung disease later in life. One study measured the workplace exposure of tunnel construction workers. Part of the study compared 115 drill and blast workers with 220 outdoor concrete workers. Total dust exposure was measured in milligram years per cubic meter (mg.y/m$^3$). The mean exposure for the drill and blast workers was 18.0 mg.y/m$^3$ with a standard deviation of 7.8 mg.y/m$^3$. For the outdoor concrete workers, the corresponding values were 6.5 mg.y/m$^3$ and 3.4 mg.y/m$^3$.

(a) The sample included all workers for a tunnel construction company who received medical examinations as part of routine health checkups. Discuss the extent to which you think these results apply to other types of workers.

(b) Use a 95% confidence interval to describe the difference in the exposures. Write a sentence that gives the interval and provides the meaning of the 95% confidence.

\[
\begin{align*}
\bar{x}_{db} &= 18.0 \text{mg.y/m}^3, \quad s_{db} = 7.8 \text{mg.y/m}^3, \\
\bar{x}_{cw} &= 6.5 \text{mg.y/m}^3, \quad s_{cw} = 3.4 \text{mg.y/m}^3.
\end{align*}
\]

\[
18.0 - 6.5 \pm \frac{1.981 \sqrt{7.8^2 + 3.4^2}}{\sqrt{115 + 220}}
\]

(9.99, 13.01)

Now for the meaning of our interval, which is the most important part. We found that the average difference in dust exposure of outdoor workers compared to the drill and blast workers could be between 9.99 mg.y/m$^3$ to 13.01 mg.y/m$^3$, with 95% confidence in our statement. The blast and drill workers being on average exposed to more than the outdoor workers.

(c) Test the null hypothesis that the exposures for these two types of are the same. Justify your choice of a one-sided or two-sided alternative. Report the test statistic, the degrees of freedom, and the P-value. Give a short summary of your conclusion.

I will use a two-sided test, since while unlikely, it may be the case that drill and blast workers use protective equipment that would in fact limit the exposure to dust compared to an outdoor worker.

\[
t = \frac{18.0 - 6.5 - 0}{\sqrt{7.8^2 + 3.4^2 \over 115 + 220}} = 15.07
\]

\[
P(t > 15.07) \approx 0
\]

The result clearly indicates that if we assumed that there is no difference in average exposure between the two types of worker the likelihood of seeing a difference as large as the one observed would be almost impossible to observe by chance alone. Thus, we have enough evidence to reject the null hypothesis.
(d) The authors of the article describing these results note that the distributions are somewhat skewed. Do you think that this fact makes your analysis invalid? Give reasons for your answer.

Again the sample sizes are large for both and almost equal in size which is very important when comparing two means. Since it is mentioned that the distributions are somewhat skewed, meaning no extreme outliers, then the results should be valid.

7.57 A recent study of food portion sizes reported that over a 17-year period, the average size of a soft drink consumed by Americans aged 2 years and older increased from 13.1 ounces (oz) to 19.9 oz. The authors state that the difference is statistically significant with \( P < 0.01 \). Explain what additional information you would need to compute a confidence interval for the increase, and outline the procedure that you would use for the computations. Do you think that a confidence interval would provide useful additional information? Explain why or why not.

7.75 (e) Find a 95% confidence interval for the difference in mean DBHs? Explain how this interval provides additional information about this problem.

\[
34.5 - 23.7 \pm 2.045 \sqrt{\frac{306.3}{30} + \frac{203.3}{30}}
\]

(2.4cm, 19.3 cm)

t-Test: Two-Sample Assuming Unequal Variances

<table>
<thead>
<tr>
<th></th>
<th>DBH Northern</th>
<th>DBH Southern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>23.7</td>
<td>34.5333333333</td>
</tr>
<tr>
<td>Variance</td>
<td>306.2551724</td>
<td>203.2995402</td>
</tr>
<tr>
<td>Observations</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>-2.628617061</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.005522342</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.672522103</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.011044684</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.003239388</td>
<td></td>
</tr>
</tbody>
</table>