Chapter 7 HW Solutions

7.63  
  a) The null hypothesis should involve \( \mu_1 \) and \( \mu_2 \), not \( \overline{x}_1 \) and \( \overline{x}_2 \).
  
  b) The 56 males of the study are counted both in population 1 and in population 2. There should be complete independence between these populations.
  
  c) A \( P \)-value of 0.94 has no meaning. Remember that the only purpose of a \( P \)-value is to assess the weight of the evidence against the null hypothesis and in favor of the alternative.
  
  d) The alternative \( \mu_1 < \mu_2 \) is equivalent to \( \mu_1 - \mu_2 < 0 \). But the data gave a positive \( t \)-statistic, meaning that \( \overline{x}_1 - \overline{x}_2 > 0 \)! Thus the data does not favor the alternative hypothesis, and thus no \( P \)-value should be reported.

7.64  
  a) Remember that we reject the null hypothesis \( \mu_1 - \mu_2 = 0 \) against a two sided alternative at the \( \alpha \) level precisely when 0 does not lie in the \( 1 - \alpha \) level confidence interval

\[
\overline{x}_1 - \overline{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}
\]

(We saw this fact in Lab 5, part a.) Since 0 does not in the confidence interval, we would reject the null hypothesis.
  
  b) Larger samples generally give smaller margins of error.

7.66  
For this we just use the formula given on the worksheet:

\[
\overline{x}_1 - \overline{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = -15 \pm 2.042 \sqrt{\frac{19^2}{50} + \frac{16^2}{40}}
\]

\[
= -15 \pm 7.536
\]

This interval includes fewer values than would be included with a 99% confidence interval.

7.68  
  a) A 95% confidence interval for class prep is

\[
706 \pm 1.962 \cdot \frac{526}{\sqrt{1839}}.
\]

  b) A 95% confidence interval for number of minutes per week on Facebook is

\[
742 \pm 1.962 \cdot \frac{651}{\sqrt{1839}}.
\]

Note that we got these numbers by multiplying the given numbers by 7 (since the given numbers are in terms of days, and we want weeks.)
  
  c) These populations might be skewed to the right, since it is quite likely that each population includes some obsessive members (i.e. people that study all the time, or people that use Facebook all the time)

7.71  
I solved this problem using R. I’ll outline my solution below. First, I load the data set, and called it data. Then, I separated the sad elements from the neutral. To do this, I used the following code:

```r
sad_index = data$Group == 'S'
sad = data$Price[sad_index]
neutral_index = data$Group == 'N'
neutral = data$Price[neutral_index]
```
Once I’d separated the sad data from the neutral data, I formed histograms:

\[ \text{hist(sad)} \]
\[ \text{hist(neutral)} \]

The results are displayed below:

a) The histograms are as follows:

![Histogram of sad](image1)

![Histogram of neutral](image2)

b) The table is as follows:

<table>
<thead>
<tr>
<th></th>
<th>sad samp</th>
<th>neutral samp</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>( \bar{x} )</td>
<td>2.12</td>
<td>0.57</td>
</tr>
<tr>
<td>( s )</td>
<td>1.24</td>
<td>0.73</td>
</tr>
</tbody>
</table>

c) \( H_0 \) : \( \mu_{\text{sad}} = \mu_{\text{happy}} \) and \( H_a \) : \( \mu_{\text{sad}} > \mu_{\text{happy}} \), where \( \mu_{\text{sad}} \) is the mean spending of the sad population and \( \mu_{\text{happy}} \) is the mean spending of the happy population.

d) There are 13 degrees of freedom for this test, the \( t \)-statistic is 4.30, and the \( P \)-value for this data is about .0004. We reject \( H_0 \) at the 5% level of significance.

You can run the t-test in R as follows:

\[ \text{t.test(sad,neutral,alternative="greater")} \]

Note the results are close to those obtained above, but not exact: we will discuss the differences in class later on.

e) A 95% confidence interval would be

\[
1.55 \pm 2.16 \cdot \sqrt{\frac{1.55}{17} + \frac{.53}{14}} = 1.55 \pm .776
\]