Exercise Set #30
Answers to Odd-Numbered Exercises

30-1 (a) (v) and (vi)

(b) Since the hypothesis test concerns the possible relationship between two qualitative variables, a chi-square test concerning independence should be performed.

(c) 

\begin{itemize}
  \item \textbf{Step 1} \quad \text{H}_0: \text{ Hair color and eye color are independent. } \\
  \text{H}_1: \text{ There is a relationship between hair color and eye color. } \\
  \alpha = 0.05 \\
  \text{OR} \\
  \chi^2 = 10.929
  \\
  \text{Step 2} \\
  \chi^2_1 = 10.929 \\
  \text{Step 3} \\
  3.841 = \chi^2_{1; 0.05} \\
  \text{reject H}_0 \quad p \text{-value} < 0.001
  \\
  \text{Step 4} \\
  \text{Since } \chi^2_1 = 10.929 \text{ and } \chi^2_{1; 0.05} = 3.841, \text{ we have sufficient evidence to reject } H_0. \text{ We conclude that there is a relationship between hair color and eye color (}p\text{-value} < 0.001\text{). Since } H_0 \text{ is rejected, we need to describe the relationship.}
\end{itemize}
(d) Either of the answers below is correct (and two bar or pie charts is also correct).

<table>
<thead>
<tr>
<th>Hair Color</th>
<th>Eye Color</th>
<th>Light</th>
<th>Dark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light</td>
<td>39.7%</td>
<td>60.3%</td>
<td>100%</td>
</tr>
<tr>
<td>Dark</td>
<td>66.7%</td>
<td>33.3%</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hair Color</th>
<th>Eye Color</th>
<th>Light</th>
<th>Dark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light</td>
<td>29.4%</td>
<td>55.9%</td>
<td></td>
</tr>
<tr>
<td>Dark</td>
<td>70.6%</td>
<td>44.1%</td>
<td></td>
</tr>
</tbody>
</table>

It appears that the proportion of individuals with light eyes is higher among individuals with dark hair than among individuals with light hair.

(e) The expected frequencies 35, 28, 50, and 40, are all greater than 5.

(f) Since $H_0$ is rejected, the Type I error is possible, which is concluding that a relationship exists when actually hair color and eye color are independent.

(g) $H_0$ would have been rejected with $\alpha = 0.01$ and with $\alpha = 0.10$. 
(b) Since the hypothesis test concerns the possible relationship between two qualitative variables, a chi-square test concerning independence should be performed.

(c) Since $\chi^2_3 = 60.885$ and $\chi^2_3; 0.05 = 7.815$, we have sufficient evidence to reject $H_0$. We conclude that there is a relationship between sex of the student and type of major selected ($p$-value < 0.001). Since $H_0$ is rejected, we need to describe the relationship.
(d) It appears that the proportion of mathematical science majors and the proportion of natural science majors are both higher among males than among females.

(e) All expected frequencies are greater than 5.

(f) Since $H_0$ is rejected, the Type I error is possible, which is concluding that a relationship exists when actually sex of the student and type of major selected are independent.

(g) $H_0$ would have been rejected with $\alpha = 0.01$ and with $\alpha = 0.10$. 

<table>
<thead>
<tr>
<th>Sex</th>
<th>Mathematical Sciences</th>
<th>Natural Sciences</th>
<th>Social Sciences</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>45.25%</td>
<td>46.25%</td>
<td>6.75%</td>
<td>1.75%</td>
</tr>
<tr>
<td>Female</td>
<td>35.75%</td>
<td>34.00%</td>
<td>23.00%</td>
<td>7.25%</td>
</tr>
</tbody>
</table>
30-5 (a) Since the data consist of two independent random samples, the separate two sample \( t \) test should be performed.

(b) Step 1
- \( H_0: \mu_N - \mu_H = 0 \)
- \( H_1: \mu_N - \mu_H > 0 \)
- \( \alpha = 0.05 \) (one-sided)

Step 2
- \( n_N = 8, \quad \bar{x}_N = 1.37 \) inches, \( s_N = 0.0918 \) inches,
- \( n_H = 11, \quad \bar{x}_H = 1.30 \) inches, \( s_H = 0.1104 \) inches,
- \( t_{17} = 1.506 \)

Step 3
- do not reject \( H_0 \)
- \( 0.05 < p\text{-value} < 0.10 \)

Step 4
Since \( t_{17} = 1.506 \) and \( t_{17; 0.025} = 1.740 \), we do not have sufficient evidence to reject \( H_0 \). We conclude that the mean length of ball bearings produced by the Nuketown factory is not larger than for those produced by the Highville factory \( (0.05 < p\text{-value} < 0.10) \).

(c) A confidence interval for the difference in mean length of ball bearings produced by the Nuketown and Highville factories would not be of interest, since the null hypothesis was not rejected which suggests there is no difference in mean length of ball bearings produced by the Nuketown and Highville factories.
(d) five-number summary for Nuketown:  
   1.21, 1.305, 1.395, 1.43, 1.49  
five-number summary for Highville:  
   1.16, 1.22, 1.28, 1.37, 1.55  

Since there are no potential outliers, and each distribution looks only slightly to moderately skewed, the $t$ statistic could be considered appropriate.

(e) Since $H_0$ is not rejected, the Type II error is possible, which is concluding that $\mu_N - \mu_H = 0$ when actually $\mu_N - \mu_H > 0$.

(f) $H_0$ would not have been rejected with $\alpha = 0.01$ but would have been rejected with $\alpha = 0.10$. 
30-7 (a) A chi-square test concerning independence requires observations on two qualitative variables, but this data will consist of observations of one qualitative variable (political party).

(b) A chi-square goodness-of-fit test should be considered with one random sample of observations of a qualitative variable (political party).

(c) In order that the expected frequency for each of the four categories be at least 5, the sample size must be at least 20.