One of the questions asked of a random sample of adult Americans on the 2004 General Social Survey was:

From time to time, most people discuss important matters with other people. Looking back over the last six months - who are the people with whom you discussed matters important to you? Just tell me their first names or initials.

The interviewer then recorded how many names each person gave, with the person’s gender.

a) The relevant parameter for this study can be symbolized as $\mu_{\text{men}} - \mu_{\text{women}}$. Describe what this parameter means in this context.

b) State the appropriate null and alternative hypotheses (in symbols) for testing whether American men and women differ with regard to average number of close friends.

The survey responses are summarized in the following table (and in the datafile CloseFriends.txt, available on the HW page of our course website):

<table>
<thead>
<tr>
<th>Number of close friends</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of men responses</td>
<td>196</td>
<td>135</td>
<td>108</td>
<td>100</td>
<td>42</td>
<td>40</td>
<td>33</td>
<td>654</td>
</tr>
<tr>
<td>Number of women responses</td>
<td>201</td>
<td>146</td>
<td>155</td>
<td>132</td>
<td>86</td>
<td>56</td>
<td>37</td>
<td>813</td>
</tr>
</tbody>
</table>

c) Use R to produce histograms for comparing the distribution of number of close friends between men and women. Comment on what the histograms reveal about the shapes of the distributions.

d) Use R to determine the sample mean and sample standard deviation of the number of close friends for each sex. Report these with appropriate symbols. Also show how to calculate the sample means by hand from the table above.

e) Conduct a two-sample $t$-test of the hypotheses from a). Show how to calculate the test statistic by hand. Report the test statistic and $p$-value. State your test decision at the .05 significance level, and summarize your conclusion.

f) Produce a 95% confidence interval for the difference in population means (for the number of close friends) between men and women. Show how to calculate the confidence interval by hand. Also write a sentence or two interpreting what the interval reveals.
g) Are the technical conditions for the two-sample $t$-test satisfied here? Explain.

h) Now conduct a test of whether these sample data suggest that the proportion of Americans who say they have zero close friends differs between men and women. Report the hypotheses, test statistic, and p-value. State your test decision at the .05 significance level, and summarize your conclusion.