1. Refer to the following generated data from a hypothetical study of classroom interventions, relating to whether the students passed (“yes”) or failed (“no”) a particular assessment:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>168</td>
<td>43</td>
</tr>
<tr>
<td>Control</td>
<td>359</td>
<td>146</td>
</tr>
</tbody>
</table>

a. What is the Risk of failing the assessment for those receiving the intervention?

b. What is the Risk of failing the assessment for those not receiving the intervention?

c. What is the Relative Risk of failing for the Intervention group versus the Control group? (Careful, this one can be tricky.)

d. What can you say about the Relative Risk in this example? In other words, interpret the results from Part 1C, making sure to reference the numerical results in a meaningful manner. Also, make sure to answer whether the intervention is effective or not.

2. Again, refer to the data in Question 1.

a. What is the odds ratio of failing the assessment for those who received the intervention versus for those who did not receive the intervention?

b. What is the odds ratio of failing the assessment for those who did not receive the intervention versus for those who did receive it?

c. Is the Odds Ratio significant? What is the p-value? What is the confidence interval for each comparison?

d. What can you say about the odds of failing in this example? In other words, interpret the results from Part 2C, making sure to reference the numerical results in a meaningful manner. Also, make sure to answer whether the intervention is effective or not.

3. R calculated the ratio confidence intervals (for risk and for odds) for you. How would you do each by hand?

**Hint:** Remember the discussion during lecture about the log of the odds and taking the exponent (which is the opposite of a log).

**Note:** Show your calculations if you would like to illustrate your point or check your work, but your answer should be a few sentences describing the process.

4. In the United States, the estimated annual probability that a woman over the age of 35 dies of lung cancer equals 0.001304 for current smokers and 0.000121 for nonsmokers. (Question 2.7 in the Agresti book, pg. 56).

a. Calculate and interpret the difference of proportions and the relative risk. Which is more informative for these data? Why?

b. Calculate and interpret the odds ratio. Explain why the relative risk and odds ratio take similar values.