Statistics 101: Section L – Laboratory 11

In lab last week you looked at the eye color of students in a small population (250 students). Consider a much larger population (millions of people) where the population proportion of blue eyes is 0.312. In the first activity, you will look at simulating drawing samples from this population at random and investigate how different choices of sample size and confidence level affect the confidence intervals constructed from these samples. In the second activity you will investigate what affect violating the success/failure condition has on the confidence level.

You will use the applet on the web at

http://www.rossmanchance.com/applets/Confsim/Confsim.html

You can get to this URL from the course web page via the link to Confidence Interval for a Population Proportion under Computing. In the applet, the method should be Proportions – Wald. Enter 0.312 for the value of $\theta$.

Activity 1

a) Enter the value of 10 for $n$ (a random sample of size 10) and leave the number of Intervals at 1, then click on Sample. Right click on the interval produced and record the lower limit, center and upper limit of this confidence interval. Does the interval displayed contain 0.312?

b) Change the value of $n$ to 25 and press the sample button. How does the interval change? Again record the lower limit, center and upper limit of this confidence interval.

c) Change the value of $n$ to 100 and press the sample button. How does the interval change? Again record the lower limit, center and upper limit of this confidence interval.

d) Why is the center of the interval different for a), b) and c)?

e) How does the length of the interval (upper limit – lower limit) change as you increase the sample size?

f) Enter the value of 10 for $n$, leave the number of Intervals at 1, and click on Sample. Record the lower limit, center and upper limit of this confidence interval.

g) Change the confidence to 80% and press the recalculate button. How does the interval change? Again record the lower limit, center and upper limit of this interval.

h) Change the confidence to 99% and press the recalculate button. How does the interval change? Again record the lower limit, center and upper limit of this interval.

i) How does the length of the interval (upper limit – lower limit) change as you change the level of confidence?

j) A narrow confidence interval is desirable. Give two things that you can do to get a narrower confidence interval. Which of these two ways do you think is better? Explain briefly.
Activity 2

a) Start with $\theta = 0.312$, $n = 10$, intervals = 250 and conf. level = 95%. Press the sample button. Report the value for the intervals containing 0.312.

b) Press the sample button again and report the value for the intervals containing $p$ and the running total. The running total combines the results of the first 250 samples in a) with those of the second 250 samples in b).

c) Click the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

d) Reset the simulation and change the conf. level to 90%. Click the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

e) Is the success/failure condition satisfied when $p = 0.312$ and $n = 10$?

f) Reset the simulation and start with $\theta = 0.312$, $n = 25$, intervals = 250 and conf. level = 95%. Press the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

g) Reset the simulation and change the conf. level to 90%. Click the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

h) Is the success/failure condition satisfied when $p = 0.312$ and $n = 25$?

i) Reset the simulation and start with $\theta = 0.312$, $n = 100$, intervals = 250 and conf. level = 95%. Press the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

j) Reset the simulation and change the conf. level to 90%. Click the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

k) Is the success/failure condition satisfied when $p = 0.312$ and $n = 100$?

l) Reset the simulation and start with $\theta = 0.312$, $n = 400$, intervals = 250 and conf. level = 95%. Press the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

m) Reset the simulation and change the conf. level to 90%. Click the sample button enough times so that the running total summarizes 10,000 samples. Report the running total. What should this running total be?

n) Is the success/failure condition satisfied when $p = 0.312$ and $n = 400$?

o) What can you conclude about the success/failure condition and the confidence one has in the procedure that produces confidence intervals?
### Activity 1:

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Lower Limit</th>
<th>Center</th>
<th>Upper Limit</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) n = 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) n = 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) n = 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d) Why are the centers different?

e) How do the lengths change?

<table>
<thead>
<tr>
<th>Conf. Level</th>
<th>Lower Limit</th>
<th>Center</th>
<th>Upper Limit</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>f) 95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g) 80%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h) 99%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

i) How do the lengths change?

j) Two ways to get a narrower interval. Which is better?
Activity 2

Start with $\theta = 0.312$, $n = 10$, intervals = 250 and conf. level = 95%.

a) Report the value for the intervals containing $p = 0.312$.

b) Report the value for the intervals containing $p$ and the running total.

c) – n)

<table>
<thead>
<tr>
<th></th>
<th>95%</th>
<th>90%</th>
<th>success: $np &gt; 10$?</th>
<th>failure: $n(1-p) &gt; 10$?</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 10, p = 0.312$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 25, p = 0.312$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 100, p = 0.312$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 400, p = 0.312$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

p) What can you conclude about the success/failure condition and the confidence one has in the procedure that produces confidence intervals?