MONT 107N – Understanding Randomness
Group Discussion 2
February 22, 2010

Background and Goals

We have seen how to estimate percentages by sampling a population, using the insight provided by 0/1 box models. Because of the Central Limit Theorem, we can use the normal curve to estimate the chances of certain results of the sampling process. The goal for today is to practice on several problems of this type, in order to "solidify" our understanding of the ideas and processes involved.

Discussion Questions

A. The point of this question is to reinforce the idea that in finding estimates of percentages by sampling, we actually use *an estimate, or approximation* to the standard error – this leads to what Freedman, Pisani, and Purves call the "bootstrap." Suppose there is a 0/1 box with 10,000 tickets, and exactly 2000 of them are 1's.

1. What is the actual exact value of the standard error for 400 draws from the box?
2. Now a person (who does not know the true percentage of 1's in the box) draws a random sample of 400 tickets, and finds 75 1's in that sample. How would that person estimate the percentage of tickets in the box and the standard error?
3. Same question as in part 2, but for a second person who draws a different random sample containing 87 1's.
4. Is the true value of the percentage of 1's within 2 (estimated) SE's of the estimate of the percentage in parts 2 and 3?

B.
1. A random sample of size 400 was taken from the population of all manufacturing companies in a big state. Of the 400, 11 had more than 100 employees. Estimate the percentage of manufacturing companies with more than 100 employees.
2. In the situtation of part 1. estimate the standard error for this percentage estimate.
3. How likely is it that the true percentage of companies with more than 100 employees was greater than 4%?
4. In the same state as in the previous parts, a random sample of 400 was taken from the population of all individuals employed by manufacturing companies. A total of 187 individuals in this sample worked for companies with more than 100 employees. Estimate the percentage of of workers employed by companies with more than 100 employees.
5. Estimate the standard error for the percentage from part 4.
6. What is the chance that fewer than 40% of the individuals employed by manufacturing companies are employed by companies with more than 100 employees?
7. Is the difference between your answers to parts 1 and 4 due to chance error?