HW 2

September 7, 2021

1. What is a simple random sample? If all sampling units have the same probability of being in the sample, is this always a simple random sample? Why or why not?

2. Use R to take a simple random sample of size n = 4 (assume without replacement) from this population: 10, 11, 13, 11, 10, 6, 22, 15, 14, 23. Please include your output in the homework. Using this sample, compute your estimator of the population mean and find its standard error (what IS a standard error?) and a 95 percent confidence interval.

3. We want to know the total number of moose in an area. Suppose we have divided the region into $N = 200$ quadrats, our guess is that the standard deviation of the moose counts is around $s = 3$ moose AND we would like a margin of error of less than plus or minus 100 moose. How many sampling units do we have to visit?

4. Finally (see problem 4), suppose we had decided to sample $n = 12$ of the quadrats and got a sample average of 14 moose per quadrat and a sample variance of $s^2 = 125$ square moose per quadrat. Find an estimate of the total number of moose, along with its standard error and then construct a 95 percent confidence interval for the total number of moose.

5. I wish to estimate the total number of squirrels in a large region. I'll do that by dividing the region into $N = 1500$ transects (long, narrow plots, with one end against a road), each 10 m wide and 1 Km long. I select $n = 10$ of these to visit (I'll walk the transects and count animals. I'll assume that I count every one in the transect and ignore all animals outside the transect. This might work if I alarm them as I go by.) I get the following counts: $12, 20, 8, 42, 23, 18, 6, 8, 13, 17$.

(a) Find an estimator of the total number of squirrels in the entire region and the standard error, along with a 95 percent confidence interval.

(b) If I divide the estimated total by the total area of the region (in square Kilometers) I'll get a density in squirrels per Km^2. Find a 95 percent confidence interval for this density (hint - how do you adjust the standard error from part a?).

6. To use the typical estimator ± 2 standard errors to find a 95 percent interval for a proportion, you need the sampling distribution of the sample proportion to be close to normal. You can assume this is the case if $n * (\text{est proportion}) > 10$ and $n * (1 - \text{est proportion}) > 10$.

Suppose we are looking for the proportion of spruce trees in a low land forest that have a certain genetic trait. We somehow make a list of N = 1320 trees in the area. We take a SRS of size n = 120 trees and find that 13 have the genetic trait.

(a) Is this sample size sufficient for us to assume the sampling distribution of p.hat is approximately normal? Why or why not?

(b) Find a 95 percent confidence interval for the true proportion of trees in the region with the trait.

7. I wish to conduct a political poll in a small town. The town has a total of 450 residents and I can actually take a SRS of them. I would really like a margin of error of, at most, plus or minus 5 percent. What sample size do I need to take? What sample size would I need to take if I perversely decided to use SRS WITH replacement?

8. (This will require the use of some R programming.) Use bootstrapping on the data in problem 2 to get a 95 percent confidence interval for the population mean.

For problem 9, consider the following data, which are the actual number of trees in each of the N = 70 plots in the region:

1	3	4	5	3	3	0	7	5	0
1	4	1	4	2	3	4	3	4	4
3	3	0	1	2	1	5	5	3	3
11	9	9	15	10	6	6	12	12	12
18	14	13	7	10	13	3	14	11	11
15	11	13	12	12	13	19	10	11	14
13	6	9	7	15	14	9	16	13	12

9. Take a SRS (without replacement) of size n = 18 from the population in the table. Find a 95 percent confidence interval for the total over the entire area. Later you will compare this estiamte with the estimate from a stratified random sample.