Estimating Populations: Mark-Recapture Sampling

Overview

This exercise is a simulation of a classic technique used for estimating the size of animal populations. Using dried beans to represent individual organisms, students will randomly select and then mark a sample from the population. After these marked individuals are mixed back in with the rest of the population, students take another sample, making note of the number of marked individuals "caught" a second time. The mathematical relationship used to estimate the total number of individuals in the original population is based on setting up a proportion. By increasing sample sizes and number of trials, students will experience basic statistical concepts.

North Carolina Standard Course of Study

Competency goal 7: The learner will conduct investigations and use technologies and information systems to build an understanding of population dynamics.

Textbook References

McDougal Littell

Unit D, chapter 2.1 (pp 45-53) explores populations. This exercise replaces a mark-and-recapture exercise on pages 52-53.

Prentice Hall

Chapter 11, section 2 (pp 391-401) introduces populations. Mark-and-recapture is addressed on page 393.

Background

Biologists often need to count organisms. They may want to count cells on a microscope slide in order to test for disease. They may want to count trees in a forest in order to assess the health of an ecosystem. Or they may want to count salamanders in a stream in order to assess the impact of pollution. Biologists use different techniques for each of these measurements. For many kinds of organisms, it is virtually impossible to count every single individual. And if biologists want to count organisms that can move around (immigrate or emigrate), reproduce, or die, then they have to use methods that take these kinds of fluctuations into account.

One sampling method used for estimating animal populations is known as *mark-recapture sampling* or sometimes as *capture-recapture sampling*. For example, a biologist might set live traps for a certain kind of beetle. Once collected, each beetle could be marked with a dab of fingernail polish and then be released back to the wild. After a certain amount of time, say a week, the traps are reset and another sample of beetles is caught. Some of the beetles in this second group will be new ones, never seen before, but some of them may be marked individuals from the first trapping. To keep things straight, let's assign some letters to each of these numbers:

Table 1	
N	The ***total (N)umber*** of beetles in the whole population (the number we are looking for)
M	The total number of ***(M)arked*** beetles from the *first* trapping
n	The ***(n)umber*** of beetles in the *second* sample (both marked *and* unmarked)
R	The number of ***(R)ecaptures*** (that is, the number of marked beetles in the second sample)

One of the conventions often used by scientists is to label a whole population with an uppercase letter (**N**) and a subset of that population with a lowercase letter (**n**).

Using the *parameters* we have just defined, our beetle biologist sets up the following *proportion*:

$$\frac{\mathbf{N}}{\mathbf{M}} = \frac{\mathbf{n}}{\mathbf{R}}$$

This proportion says that the *ratio* of the ***total*** (N) number of beetles to the total number of ***marked*** (M) beetles in the first sample is equal to the *ratio* of the total (n) number of beetles in the second sample to the number of ***marked (recaptured)*** (R) beetles in the second sample. By doing a very little bit of algebra, multiplying both sides of this equation by **M**, we get:

$$\mathbf{N} = \frac{\mathbf{Mn}}{\mathbf{R}}$$

Since the biologist knows the actual numbers for each of the variables on the right side of this equation, he or she can simply 'plug in' those numbers and get a value for the total population.

This mark-recapture technique makes certain *assumptions*. Making assumptions is an essential step in all scientific research. Perhaps the most fundamental assumption scientists make is that their observations on particular individuals or groups of organisms apply in general to the rest of the organisms that they never see. Thus, the assumptions must be carefully thought out in hopes that they reflect reality. It is important to state assumptions explicitly, because two scientists with different assumptions may interpret the exact same data in different ways. The following are some of the assumptions used in mark-recapture sampling:

1. First, we assume that every beetle in the population has an equal chance of being captured. Thus, both of our samples must be *random* samples.
2. Secondly, we assume that there is no change in the ratio between marked and unmarked beetles during the interval between samplings. This means, for example, that the marking technique does not make a beetle more likely to be eaten by a predator. If that were the case, then after a week all the marked beetles might be eaten and therefore removed from the population. It is actually okay if some of the beetles in the population emigrate or die, just so long as there is no difference between marked and unmarked animals in these regards.
3. A third assumption is that the marked animals, when they are released back into the wild, distribute themselves randomly. If all the marked beetles were to remain near the trap (because the fingernail polish glued their wings together, for example), then they might be more likely to be caught in the second sample. This would upset the ratio that we are hoping to preserve.
4. Finally, we must assume that there is not a significant increase in the size of the population due to births or immigration of unmarked beetles. This also would change the ratio of marked to unmarked beetles.

In other words, the samplings must be random, and the time between samplings must be long enough to allow for thorough mixing of marked animals, but not so long as to allow for a significant increase by immigration or reproduction.

It turns out that, for technical reasons, the equation above actually tends to over-estimate the size of the total population. There are ways to fix this by making changes to the equation (by adding and subtracting some 1s in judicious places), but these fixes are not important for our purposes. On the other hand, the larger the number of *recaptures*, the better the estimate of the total population. Thus, it is important to get a large second sample (**n**) in order to assure a large recapture (**R**).

The following table displays data collected in a typical run. All trials started with 250 beans. Samples were made with a 1 ounce paper cup. Note that the range of values of the estimated population decreases with increased sample size (from 483 to 90 to 83). Also note that the estimates get closer to the real population size as the sample size increases (78 above, 26 below, 9 above). In this example, when the samples were about one fourth of the real population (69 out of 250), the estimate was very close (259). Finally, note what happens in the fifth row of the table when no marked beans are in the second sample. This type of serious error becomes less likely when the sample sizes are larger.

Sample size	First Sample Marked (M)	Second Sample Total (n)	Second Sample Marked (R)	Calculated Population Estimate (N)	Actual Population
1 oz	23	25	3	192	250
1 oz	23	21	1	483	250
1 oz	23	21	1	483	250
1 oz	23	21	1	483	250
1 oz	23	25	0	0	250

Range of estimates	0 to 483 = 483
Average of estimates	328
Actual - Average	78

2 oz	45	47	8	264	250
2 oz	45	40	10	180	250
2 oz	45	43	9	215	250
2 oz	45	42	7	270	250
2 oz	45	46	11	188	250

Range of estimates	180 to 270 = 90
Average of estimates	224
Actual - Average	-26

3 oz	69	73	19	265	250
3 oz	69	62	14	306	250
3 oz	69	67	17	272	250
3 oz	69	68	21	223	250
3 oz	69	73	22	229	250

Range of estimates	223 to 306 = 83
Average of estimates	259
Actual - Average	9

Materials

Materials for the whole class

- Bag of large lima beans

Materials for each pair of students

- Colored pencils
- Paper bag
- 1-oz. cup for measuring out beans

Preparation

- Be sure all students have a science notebook.
- Students will work in pairs.
- Students will do some math calculations; calculators might be helpful.

Procedure

<u>Part I</u>

- Ask students, "If you were a biologist studying the robins in your area, how might you determine the size of their population?
    - After students share their ideas, explain that this is a real issue that scientists who study animals and their populations often have to figure out.
    - Inform students that this activity simulates how scientists accurately estimate populations in nature.
- Instruct students to work in pairs and hand out the supplies.
- Tell students to construct a data table in their notebooks that will collect the following data: Probably one row, five columns.
    - First Sample--Marked
    - Second Sample--Total
    - Second Sample--Marked
    - Calculated Population Estimate
    - Actual Population

- Their table might look like this:

First Sample Marked	Second Sample Total	Second Sample Marked	Calculated Estimate	Actual Population

- Have each pair of students get their "population" of beans by having them grab four or five handfuls of beans from the class supply bag and put them in their paper bag.
- Instruct each pair to scoop 1 ounce (1 cup) of beans from their bag.
- Using a colored pencil, students should mark *both sides* of these beans, top and bottom, with a *line*. They should then count these marked beans and record this number in the *First Sample, Marked* column of the data sheet.
- Students should now return the marked beans back into the population. Tell them to shake the bag so the beans are mixed thoroughly.
- Again each pair should scoop 1 ounce (1 cup) of beans from their population.
- They should then count and record all of these beans in the *Second Sample, Total* column of their data sheet.
- Finally, students should also count the marked beans in the second scoop and record this number in the *Second Sample, Marked* column of their data sheet. All these beans should then be returned to the bag.
- Using the formula below, each pair of students should calculate the *Population Estimate* and record this number in their notebook.

$$\text{Population Estimate} = \frac{(\text{First Sample, Marked}) \times (\text{Second Sample, Total})}{(\text{Second Sample, Marked})}$$

- After the students have calculated their population estimate, they will have to count all the beans in their bag to determine the *Actual Population* of beans. Create three columns on your board similar to these:

Estimate (below actual)	Actual population	Estimate (above actual)

  Have each pair record their actual population of beans in the middle column and their estimate depending on whether their estimate was above or below their actual population.
- At this point, with luck, there will be a good range of values on the board. Ask students how effective they think this technique is. Those who were close and those who were not will have different responses. At this point, ask the class, "How could we make this technique more accurate? In other words, how could we narrow the range of our results?" Accept all reasonable answers. Two responses to look for are:
  - using a larger marked population
  - using the same sample size, but doing more then one trial.

Part II

- In Part II students will see what happens if they capture, mark, and recapture *more individuals*. They will also repeat this process three times to see what happens when they gather *more data over time*.
- Discuss with students their results from Part I. Remind them of their suggestions for getting better information and ask them to predict what they think will happen if they actually try some of those techniques.
- Inform the class that each group will sample beans again, but this time everyone will start with 250 beans. Using their beans from Part I, they first need to remove all of the marked beans in their sample and adjust their number to 250.
- Assign equal numbers of pairs of students to each of the following sampling groups. For a class of 30, assign 5 pairs to each group.
  - Capture and mark 2 ounces (2 cups) of beans for each trial.
  - Capture and mark 3 ounces (3 cups) of beans for each trial.
  - Capture and mark 4 ounces (4 cups) of beans for each trial.
- Each pair of students will do three capture/recapture trials, using their assigned capture size, using the following protocol. Their new data table will be like the one in Part I, but will have to have three rows, one row for each trial.
  - Remove the assigned amount of beans (2, 3, or 4 ounces).
  - Mark those beans with a colored pencil using a simple straight line on both sides of each bean.
  - Count and record the number as *First Sample--Marked*.
  - Return these beans to the whole population and mix thoroughly.
  - Again remove the assigned amount of beans (2, 3, or 4 ounces).
  - Count and record the total as *Second Sample--Total* and then count and record the marked beans as *Second Sample--Marked*. They can then calculate their estimate of the total population.
  - Now students must mark *all of the beans* in this second sample. To distinguish these beans from the previous sample, they should mark them with a line using a *different colored pencil*. For beans that were already marked with a line, they can simply add the new color to both sides.
  - Return these beans to the whole population and mix thoroughly. This will complete the **first trial**.
  - Repeat this process for the **second trial**. However, when they count the marked beans this time, they should *only count beans marked with the new color*; this includes beans with two lines, the old color and the new color. They should treat beans with only the old color as ‘unmarked.’ Teachers should check students’ data to see if their numbers are noticeably larger than for the first trial. If so, they may have included the beans with the old color.
  - After making their new population estimate, students can mark this new sample with a third color.
  - Repeat the sampling process for the **third trial**. Again, students should ignore beans that are marked only with the first two colors, and they should count only beans marked with the new color.
- Using their three trials, the students can now record the *range* and the *mean* of their estimates. And since they already know the actual population size, they can

record the *difference* between their average estimate and the actual population size, either higher or lower.

- Discuss their results.
  - Were their results closer than for Part I?
  - Were there differences depending on their sample sizes (2, 3, or 4 ounces)?
  - With the data they have collected, could they do something that might get them even better results? [They could combine the data from each of the same-sample-size groups. Does this get them even closer to the actual population size?]

Wrap-Up and Preparation for Next Class

Have all students discard any marked beans and return all unmarked beans to the bean supply bag.

Reflection/Discussion

- Lead a discussion on how this exercise relates to what scientists might do to count animals in the real world.
  - What kinds of marks would scientists use?
  - How will those marks affect the animals? [Paint might hurt them. Bright colors might make them easier for predators to find.]
  - Would scientists just catch a particular number of animals?
  - How often would they conduct their sampling?

Assessment

- Have each pair of students determine how they will present their data to show how their method in Part II compared with their original estimate from Part I.
- Present students with some data (either from another class or data that you make up yourself) and have them estimate the population.
- Present multiple sets of data of one capture size (as in Part II) and have them find the range and average. Then tell them the actual population size and have them explain how accurate the results were.
- Purposely choose data that do not provide a close estimate of the actual population. Ask them to spot any procedural errors or to explain how they could get a closer estimate.

Extensions

The following are two examples of volunteer-based animal census projects:

- The North Carolina Calling Amphibian Survey Program (CASP) is a monitoring program administered by the North Carolina Wildlife Resources Commission. For details, go to: http://www.ncparc.org/casp/casp.htm.
- The National Audubon Society and the Cornell Laboratory of Ornithology sponsor several different bird counting projects—the Great Backyard Bird Count, the Christmas Bird Count, and Project FeederWatch. For details, go to: http://www.birdsource.org/.