

Mark-Recapture Sampling

Synopsis

This exercise is a simulation of a classic technique used for estimating the size of animal populations. Using small objects such as beads or dried beans, students will randomly select and then mark a sample population. After these marked individuals are mixed back in with the rest of the objects, students again sample the population, making note of the number of marked objects "caught" a second time. The concepts of ratio and proportion are then used to estimate the total number of objects.

Objectives

After completing this exercise, students will be able to use the mathematical constructs of ratio and proportion to estimate the size of a population, and be able to describe the effects of randomness and variability on sampling methods such as this one. They will also be able to demonstrate the following abilities and understandings necessary to do scientific inquiry as outlined in the *NSES Content Standard A (Science as Inquiry)* for levels 5-8:

- use appropriate tools and techniques to gather, analyze, and interpret data,
- develop predictions and models using evidence,
- use and understand the importance of mathematics in scientific inquiry, and
- recognize that there are multiple outcomes to scientific investigations.

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. (H. B. D. Kettlewell used the technique described below in the 1950s to demonstrate natural selection in the Peppered Moth. His studies on *industrial melanism* provided the first well-documented example of the natural selection Charles Darwin invoked as the basis for his theory of evolution. This work is outlined in the **Extensions** section below.) 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 <i>first</i> trapping
n	the (n)umber of beetles in the <i>second</i> sample (both marked <i>and</i> 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*:

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

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

$$N = \frac{nM}{R}$$

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. 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, then they might be more likely to be caught in the second sample. This would upset the ratio 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**).

Procedure

Materials: For this simulation, students will need to count some objects that will take the place of real organisms. We purchased two colors of plastic beads at a craft store, but pinto beans from a grocery store will work just as well. The main criteria are that the objects be large enough to count easily, small enough that you can have lots of them, and cheap enough that they do not put a dent in the school budget. We tend to avoid objects that are edible in their countable form. Pinto beans are great. Candy, like mm's, tends to disappear during the counting process; this could cause problems with assumption #2. It is also important that the objects be 'markable.' Blue beads could replace some of the red ones and 'act' as marked beads. A dab with a permanent marker will work with a pinto bean. Black beans could 'act' as marked pinto beans, but they tend to be smaller than pintos and thus might violate two of the assumptions above. Being smaller, they might not be 'captured' at the same rate as pintos (assumption #1). In addition, being smaller, they might not mix thoroughly between the two samplings (assumption #3). Notice that the small, unpopped kernels tend to collect at the bottom of the bag of popcorn.

The only other materials needed for this simulation are a big jar to hold the total population and some small containers to hold the students' samples.

Set-up: Count out a large number of identical objects (in our example, 737 red beads), and put them in one large, see-through container. A big beaker will do nicely, or an industrial-sized mayonnaise jar from your cafeteria. It is important that there be plenty of extra room in the container to allow for thorough stirring without allowing some of the beads to 'escape' from the container and find their way under the baseboard heaters. At this point, ask the students to guess the total number of red beads (**N**). You might want to offer a prize to the student with the closest guess, but *only after the exercise is finished*. You should keep the total population size secret until the end. By the way, we prefer not to use 'clean' numbers like 1,000 or 750. Like a vacuum, nature seems to abhor clean numbers.

The next step in the set-up is to 'trap' a sizable portion of the red beads--between 15% and 20%. This should be enough to allow for good 'recapture' rates. Count these trapped beads and replace them, one for one, with beads of another color. This is the number (**M**) in the above equation, and all students will use it in their individual calculations. In our example, we replaced 124 red beads with blue ones. Mix thoroughly, in order to distribute the marked beads throughout the population (assumption #3).

Sampling: This part of the simulation needs to be done by students singly or in pairs. However, *each sample must be removed, counted, replaced, and thoroughly stirred before the next sample is taken*. This is essential, since each sample should have an equal chance of capturing any particular individual in the entire population. As a result, you will probably want your students to be doing some other activity while this is going on. Otherwise, there will be a lot of waiting around. (The exercise on **Measurement and Variability** in this series might be a good one.)

Students will record and report three numbers: (**R**) the number of marked (blue) beads, (**n**) the total size of their sample (red *and* blue together), and their estimate of (**N**) based on plugging their numbers into the equation above, using the (**M**) you have given them. Here is an example from our data:

$$N = \frac{nM}{R} = \frac{(85)(124)}{16} = 658.75 \approx 659$$

Students should round off their estimates to whole numbers. We won't have any partial beads (or beetles) in our population. (Here is a good chance to talk about the appropriateness of rounding off.)

Discussion: The numbers reported by the students should be entered in a big chart on the chalkboard or on a poster on the wall. Their estimates will almost certainly vary all over the place--some high, some low, some very high, some very close to the real number. At the bottom of the chart, the mean of all the estimates can be shown. Simply add up all the estimates and divide by the number of estimates. When we did our trial run of this exercise, we generated the numbers in Table 2.

Table 2

	# of reds in 2nd sample (R_{red})	# of blues in 2nd sample (recaptured) (R)	total in 2nd sample ($R_{red} + R$) (n)	# originally marked blue (M)	estimate of total population (nM/R) (N)
1	69	16	85	124	659
2	57	15	72	124	595
3	73	11	84	124	947
4	40	12	52	124	537
5	48	9	57	124	785
6	69	19	88	124	574
7	93	17	110	124	802
8	61	13	74	124	706
9	59	10	69	124	856
10	75	9	84	124	1157
11	44	10	54	124	670
12	61	17	78	124	569
13	54	11	65	124	733
14	63	16	79	124	612
15	43	11	54	124	609
16	72	19	91	124	594
17	37	9	46	124	634
18	63	12	75	124	775
19	86	18	104	124	716
20	59	12	71	124	734
21	71	13	84	124	801
22	57	9	66	124	909
23	52	8	60	124	930
24	39	8	47	124	729

actual population	737
----------------------	-----

mean of estimates	735
----------------------	-----

Some things to notice from this table:

- We put 737 red beads in our jar, took out 124 of them, and replaced them with blue ones--our marked beads. We did not specifically choose these numbers. There just happened to be 737 beads in the package we bought, and we grabbed our first sample so that it looked like about 20%. (It's actually about 17%.)
- We then ran 24 trials. For each trial we stirred the beads thoroughly, removed a handful, counted the reds and blues, and then returned the sample to the jar. Why 24 trials? We got tired and 24 seemed like enough repetitions. If these were beetles, we might have to be satisfied with one or a few trials. And if they were polar bears, we might be lucky to get any trials at all!
- The estimates in the right hand column varied quite a bit from the actual number shown at the bottom of the table. A few trials (#13, #20, and #24) came very close. Others (#4 and #10) were way off. Does this mean that some trials were better than others? [No. It is important to note that all of these trials were valid trials.]
- Though the individual trials varied quite a bit, the mean (average) of all those trials came *very close* to the actual number. Does this mean we did a really good job? Should we keep sampling in hopes that we will eventually get a mean that hits the actual number exactly? [No and no. Remember that

in the wild we would never know the actual number. It is nice to see that we did come so close, but that validates the model; it says little about our ability to count handfuls of beads.]

Extensions

Industrial Melanism

In the 1950s, H. B. D. Kettlewell ran a series of experiments that are now regarded as classics in the study of evolution and natural selection. (This particular account is taken from *Ecology*, by Robert E. Ricklefs, Chiron Press, 1973, but descriptions of this work can be found in most biology textbooks.) In Great Britain, there has been a long tradition among enthusiasts of making collections of butterflies and moths. In the early 1800s, these collectors would be especially pleased when they found a melanistic form of the common Peppered Moth (*Biston betularia*). The usual Peppered Moth was mostly white with little black spots (the 'pepper'). The melanistic form was just the opposite; it was mostly dark, with little white spots. And these melanistic individuals were quite rare. However, as the decades of the 1800s passed, the collectors noticed that the melanistic moths were becoming increasingly common, especially in and around cities known for their industrial development. This was, after all, the industrial revolution in Great Britain, and cities like Manchester, Birmingham, and Liverpool were thriving. By the mid-1900s, the melanistic moths made up close to 100% of the population in some of these areas and the normal (white) forms were the rare ones. On the other hand, the white forms were still the most common ones in other areas.

Through crossbreeding experiments, geneticists had determined that this melanism was an inherited trait. Therefore, Kettlewell considered the spread of the trait to be an example of evolution and an opportunity to look for evidence of natural selection. He hypothesized that something in the environment near the industrial centers was changing to give the dark moths an advantage over the light ones in both survival and reproduction. How could he test this hypothesis?

Kettlewell collected more than 3,000 caterpillars, fed them in his laboratory, and allowed them to pupate and undergo metamorphosis into adults. While he was raising the caterpillars, he selected two forest tracts--one near an industrial center, the other in a natural area. When his adult moths were ready, he marked each one with a small dab of a special paint on the underside of its wing where it couldn't be seen by predators when the moth was resting on a tree trunk. He then released his moths, waited awhile, then set traps to recapture as many as he could.

There is an important difference here between Kettlewell's goals and ours in the simulation with the beads. We were trying to estimate the total number of beads in our population. Kettlewell didn't really care about the total number of moths in his forests; he just wanted to know if there was a difference in the survival of the light and dark forms. Here are the results from one of his experiments, when he released moths near Birmingham, an industrial center:

Table 3 (polluted woods)

	white moths	dark moths
number released	201	601
number recaptured	34	205
percent recaptured	16.0%	34.1%

It looks like the dark forms survived much better than the light forms. But doesn't it make a difference that he released so many more dark moths? If he released more, it would make sense that he would recapture more. But it's the percentage that really counts here. He recaptured 34.1 % of the dark moths; that's more than twice the percentage of white moths recaptured.

Skeptics could argue that white moths were smarter and once released, they were less likely to be caught in Kettlewell's traps. Or perhaps, white moths were stronger, flew further from the release point, and thus were less likely to be recaptured. In order to account for these possibilities, Kettlewell ran a control experiment by releasing moths in a natural area, presumably unpolluted woods. Here are the results:

Table 4 (unpolluted woods)

	white moths	dark moths
number released	496	473
number recaptured	62	30
percent recaptured	12.5%	6.3%

Notice that this time, the white moths survived much better. This eliminates the possibilities of trap avoidance or differential dispersal.

These experiments confirmed for Kettlewell that the two forms of moths survived differently in the two habitats. But why? Who or what was the *agent* of natural selection? That is, who or what did the actual selecting? Kettlewell thought that in the industrial areas, soot from the coal-burning factories was settling on the tree trunks and turning them dark. As a result, the dark moths were better camouflaged in those areas and were less likely to be seen by predators, probably birds. Against normal bark, the white moths had the better camouflage. To test this hypothesis, he placed equal numbers of white and dark moths on tree trunks in natural and polluted woods. He set up a blind and simply sat and watched to see what would happen. Since Peppered Moths only fly at night, once placed on the tree trunks, they stayed right where they were. Here is what Kettlewell observed:

Table 5 (predator experiment)

	white moths eaten by birds	dark moths eaten by birds
in natural woods	26	164
in polluted woods	43	15

Clearly, many more dark moths were eaten in the natural setting, and many more white ones were eaten in the polluted woods.

Two other pieces of this story help to add the finishing touches. In the polluted woods, it turns out that it wasn't really the soot that turned the trees black, but rather it was the other pollutants from the factories that killed the lichens that normally grew on the tree trunks. White moths with black specks look remarkably like lichens. In recent years, as Britain has introduced more stringent pollution controls, the lichens are coming back to the trees, and the white moths are becoming much more common again.

Secondly, one particular bird species, the Treecreeper, ate both types of moths in equal numbers in both types of woods. This particular bird species creeps around the trunks of trees and finds its prey by seeing their silhouettes sticking out from the bark. For them, white or dark color doesn't make any difference. The exception that proves the rule!

Census 2000

[Much of this information is from: "Census Sampling Confusion," by Ivars Peterson, in *Science News*, Vol.155, No. 10, March 6, 1999, pp. 152-154.]

April 1, 2000, was Census Day in the United States. The U.S. Constitution requires that the federal government conduct a census of the nation every 10 years. This is no mean feat. In 1790, the census cost \$44,000 and found 3,900,000 people. In 1990, taxpayers forked over \$2,600,000,000 to count 248,700,000 people. (That's just over a penny apiece in 1790 and \$10.45 apiece in 1990. Is that just inflation?) Early next year, the census bureau will mail out or hand-deliver more than 90 million questionnaires, and then census workers will try to track down the people who don't return their forms. In 1990, the official census was 248,709,873 people. However, other surveys estimated that the actual number should have been closer to 253,000,000. That difference of more than 4 million is significant in that most of those people were children, racial or ethnic minorities, or people living in poverty either in inner cities or rural areas.

The official census is used in many ways, but among them are two very important purposes. First, the distribution of people determines the distribution of seats in the House of Representatives. The more populace states have more seats. Secondly, the distribution of money for federal programs like welfare is determined by where the recipients live. In particular, the groups of people who are hardest to count are also often the people who are in most need of this federal assistance. As a result of these inconsistencies, the U.S. Census Bureau has proposed a way to adjust the final figure for the 2000 census. The Bureau wants to carry out the census in the usual way and then follow up with a special survey. In randomly selected blocks across the country, census workers would go door to door and do an extra-careful count of people who were living in each house on Census Day, April 1 (no foolin'!). The main count and this special count would then be compared, looking for people who were counted in both. This is an example of a mark-recapture sampling. For the numbers in a given block:

Table 6

N the *total (N)umber* of people in the United States

M the total number of people counted--that is, (*M*)*arked*--in the *main census*

n the (*n*)**umber** of people counted in the *special count*

R the number in the *special count* who were also found in the *main census*--that is, the (*R*)**ecaptures**

From these data, the Census Bureau thinks it can get a more accurate count of the total U.S. population (**N**).

As it turns out, this is a complicated and controversial issue that might be interesting to explore with students. The outline we have presented here is a simplified explanation of the actual counting and sampling procedures. Some statisticians fear that sampling errors can be magnified by analysis techniques and the resulting figures could be further from the true figure than the unadjusted main census itself. And politicians may be worried that the extra people added to the rolls may belong to a political party other than their own. This might change their party's representation in Congress. Local government officials may or may not want their cities to show the extra inhabitants. On the one hand, a large number of homeless people might not help a city attract business investment. And on the other hand, those same people living in poverty might mean extra federal dollars for welfare.

Instructions for Students

At the end of the case, Sherlock Holmes noticed a large, earthen crock of beans on a shelf in the larder.

"Dr. Watson, to tidy up the loose ends in this case, we've got to know how many beans are in that crock?"

"Mr. Holmes, that will take hours to count."

"Watson, give me a handful to count, and you can do the rest."

Holmes soon reported that he had counted 247 beans, then returned his handful to the crock. Though he didn't tell Watson, he had also cut a tiny notch in each of his beans. While Holmes went off on another case involving a horse, a bog, and a red-bearded colonel, Watson doggedly kept counting and recounting for the rest of the day.

When Holmes finally returned at sundown, smeared with mud and smelling like a stable that had never been cleaned, Watson was in a state of frustration and perplexity.

"Holmes, I hope you are happy. I've counted and recounted these blasted beans, and I never seem to get the same number twice."

"Take heart, my good doctor. Here, mix the beans thoroughly and give me another handful."

"Holmes, what are you up to?"

"Be patient Watson. Hmm. . . ."

"I'll wager dinner, Watson, that I can tell you the answer within forty beans, though I have spent but five minutes, and you have spent five hours. Agreed?"

"That's preposterous, Holmes. I'll take that wager and spot you to brandy and cigars afterward."

"There are about 1845 beans, right?"

"Holmes! That's incredible. How did you do it?"

"Elementary, my dear Watson. You see, doing my beans twice was not as crazy as you might have thought. . . ."

* * * * *

What was Holmes up to? He was using one of the tried and true methods used by biologists when they want to census populations of organisms. It is called *mark-recapture sampling* (or sometimes *capture-recapture sampling*), and it is based on estimating a whole population by carefully counting and marking a small sample. The following simulation should help you to understand how this works.

Procedure

Your teacher will have a big container filled with lots of red beads--more than you would like to count in one sitting. Let's call that total number (**N**). This is the number you are trying to estimate. Before you do any counting, however, your teacher will do the following:

1. Remove a sample of red beads, count them, and set them aside.
2. Count out exactly the same number of blue beads.
3. Thoroughly stir these blue beads back into the large container.

It's as if the teacher marked that sample of red beads by painting them blue. Let's call this number (**M**). The teacher will tell you what this number actually is. Sherlock Holmes cut a little notch in his sample of beans to accomplish this step. If these were animals, a biologist would try to figure out a way to mark them that would not hurt them or make them conspicuous to predators. Any suggestions of what might be done to mark beetles or fish or salamanders or giraffes?

Here is the logic of how this technique works. Once the beads (or beetles or giraffes) are marked, they are returned to the original population and are stirred (or allowed to mix) thoroughly. At some later time, you (or the biologist) return and collect another sample. This time, you count your total sample, reds plus blues, (**n**), and you count the number of marked (blue) individuals in your sample (**R**). These numbers relate to each other in the following *proportion*:

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

The total population (**N**) is to the total number marked (**M**) as the total sample (**n**) is to the number recaptured (**R**).

The logic here is that if the blue beads are mixed thoroughly when they are put back, the ratio of *total* to *marked* should be the same in the second sample as in the population as a whole.

With a little algebra, we can transform the proportion above by multiplying both sides by **M**:

$$N = \frac{nM}{R}$$

Now comes your chance to estimate the total number of beads. Each student should perform the following steps:

1. Thoroughly stir all the beads in the container.
2. Remove a large handful of beads from the container.
3. Count and record the total number of beads (red + blue) in your sample (**n**).
4. Count and record the number of blue beads in your sample. These will be your 'recaptured' beads (**R**).

5. Return your sample to the original container and stir thoroughly. *Each sample must be replaced before the next student takes a sample.*

You now should know three of the four numbers in the equation above. You found n and R , and your teacher gave you M . So plug in the numbers, get a value for N , and give that number to your teacher. After everyone has had a chance to make an estimate, you can compare your results. It might be interesting to take the mean (average) of all the samples by adding them all together and dividing by the number of samples. Finally, if the teacher, like Dr. Watson, has spent time counting *all* the beads, you can compare your particular estimate and the mean to the actual number.

Do you think Sherlock Holmes got his dinner, brandy and cigar?

Copyright © 1998 by Norman Budnitz. All rights reserved.
Teachers may copy this exercise for use in their classrooms.

Revised: February 22, 2001