

Simple Random and Systematic Random Sampling

by Sophia



WHAT'S COVERED

This lesson will explain how to ensure everyone in the population has an equal chance of participating in a sample, specifically focusing on:

1. Simple Random Sample

1a. Random Number Generator

1b. Random Number Table

2. Systematic Random Sampling

1. Simple Random Sample

A **Simple Random Sample (SRS)** is a sampling method that not only ensures that everyone in the population has an equal chance of being in the sample, but also that every sample is equally likely to be the sample that's being selected.

If you've ever experienced a raffle situation, you've experienced a simple random sample. What generally happens at these events is that someone removes tickets from the raffle puts them into a bucket.

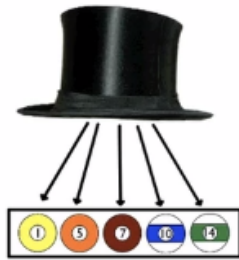
The tickets are mixed up in the bucket, and one ticket is pulled out. The owner of that ticket usually wins some kind of fantastic prize. Now, being in a simple random sample is pretty much the same thing. The only difference is that instead of winning the prize, you get to be part of the sample and that's your prize.

IN CONTEXT

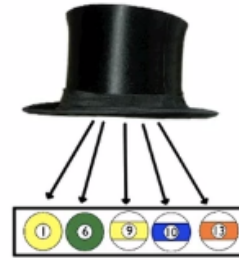
Suppose you take billiard balls from a pool table and put those all into a hat.



Next, shake it up, and pour out five billiard balls. Do this for two shakes.



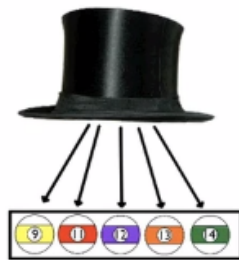
Shake #1



Shake #2

You may have noticed that the solid, yellow “1” ball was in both of these first two examples. However, it doesn't mean it's any more likely to be selected than any of the other balls. It's the same likelihood. Any sample of five, the first sample or second sample of five, were equally likely samples of five

Let's shake the hat for a third time.



Shake #3

Now, notice, all five of these were striped billiard balls, not one solid ball in the bunch. Is that unusual? Sure, it's kind of unusual to happen. Unusual samples have an equal likelihood to happen too. Just because they're strange and don't happen very often doesn't mean they can't happen. In fact, they have the same likelihood as any other selection of five.

Therefore, knowing how to take a Simple Random Sample, abbreviated SRS, is important because most inferences about the population that we do assume that we collected data in this way. So names in a hat are fine. In our case, raffle tickets in a bucket, or billiard balls in a hat...that's all fine.



TERM TO KNOW

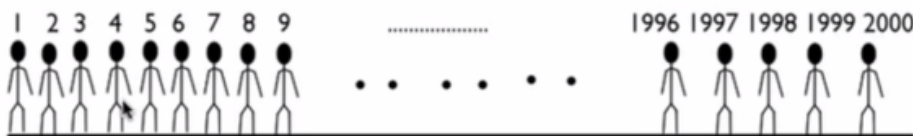
Simple Random Sample (SRS)

A method of selection that guarantees that every sample of a certain size has an equal chance of being the selected sample

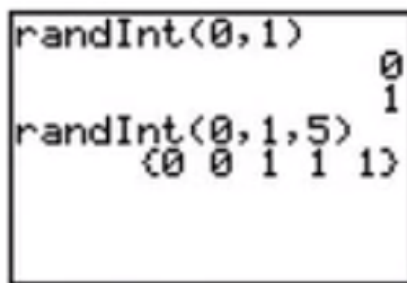
1a. Random Number Generator

However, what about the situations where we don't have the manpower to pull numbers or names from a hat? There are two other ways to take a simple random sample. One way is using a **random number generator** and the other is a random number table. First, we are going to discuss the random number generator.

⇒ **EXAMPLE** Suppose that we want to take a sample of 100 individuals from a population of 2,000 people. Below you will see some of those individuals lined up, and you can imagine that individuals 10 through 1,995 are somewhere in the middle. Each is assigned a unique number so no one can have the same number as anybody else.



Using technology such as a website, you can search "random number generator" on the internet, and websites will come up. Or, you can use a calculator. This particular model of a calculator is the Texas Instruments calculator:



randInt(1,2000,150)

"RandInt" indicates random integer--an integer is a whole number-- from 0 to 1. And so it picks either 0 or 1. When you put in the third number, it's asking how many of them do you want? In this case, you entered five. Now, you don't want numbers between 0 and 1 in this case, and you don't want five of them. You want numbers between 0 and 2,000., and you want 100 of them. Now, why was 150 written when you only want 100 numbers?

```
randInt(0,1)
0
1
randInt(0,1,5)
{0 0 1 1 1}
```

`randInt(1,2000,150)` ?

You can't select one person twice, so repeats must be ignored. It's incredibly likely that if you had just written 100 instead of 150, there would have been at least one repeat in the bunch.

Finally, you're going to select the individuals that correspond to those first 100 different numbers that were picked.

```
randInt(1,2000,150)
{8,1119,1996,732,8,125...}
```

So, person number 8, and the person that corresponds to 1,119, and the person who corresponds to 1,996 are a few that are chosen. Now, notice that the person corresponding to 8 was chosen again--you can see that it's listed twice in the list. You're not going to select that person twice because they've already been selected once, so they are crossed out. This is the reason 150 numbers were created, so you have room to cross repeats out.



TERM TO KNOW

Random Number Generator

A method of collecting a sample that utilizes technology to select random numbers corresponding to individuals in the population.

1b. Random Number Table

Using a **random number table** is basically the same idea, though it is a little bit more cumbersome. For starters, it's generally used if no technology is available. You will soon notice this is a long process and more time-consuming than using a random number generator. A random number generator typically goes faster.

Each individual is assigned a unique number, just like the random number generator; however, each member's number must have the same number of digits.

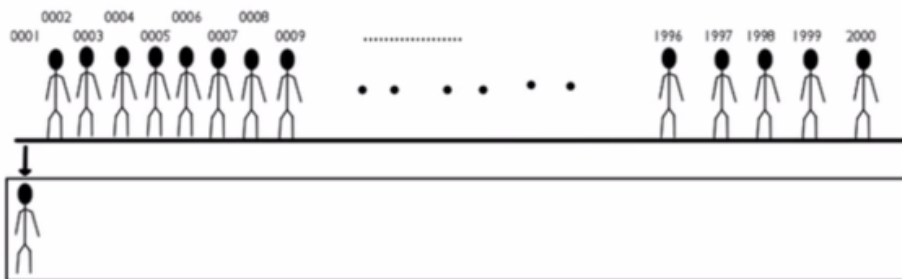
The same method as the random number generator cannot be used, because the number 2,000 has four digits, and the number 1 only has one digit. All of these must have the same number of digits, so instead of 1, it's 0001. Instead of 2, it's 0002, and so forth, all the way up to 2,000. A table of random digits can be found in a textbook or online. Four numbers will be selected at a time because each individual has four numbers.

⇒ **EXAMPLE** Suppose the first four numbers found were 1-9-2-2. That corresponds to someone in the list. There is someone who is 1,922 so that individual will be selected for the sample. It's circled in green below since a person corresponds to that number. The next number found is 3-9-5-0. No one on the list that corresponds to the number 3,950, so it is ignored. The next number, 3-4-0-5, does not correspond to an individual either, so that is ignored as well.

19223 95034 05756 28713 96409 12531 42544
 73676 47150 99400 01927 27754 42648 82425
 45467 70709 77558 00095 32863 29485 82225

You'll notice that all numbers circled in red are numbers that are unassigned in our list. This is going to make this a very cumbersome process. It will go for a while until 100 individuals are obtained. Will this work? It will work, but it might take a very long time.

One of the numbers circled in green is 0001. This is the very first person on the list, and it just happens that person 0001 will be among the sample. This individual will be selected along with everyone else whose four-digit number was selected.



TERM TO KNOW

Random Number Table

A method of collecting a sample to select random numbers corresponding to individuals in the population. Each is assigned a number, which is then selected from the table.

2. Systematic Random Sampling

There is one thing to know about systematic sampling right off the bat: it is not inherently random. You have to be very careful about this. A **systematic random sample** involves assigning a value, "k," to individuals within a population. Then, you state that every "k"th individual is chosen, similar to elementary school when you counted off by 3's to create teams.

Every 2nd person 

Every 3rd person 

Every 4th person 

The value of "k" can be anything. You could choose every second individual, in which case all the green people are in, and all these black stick figures are out. Or, you could do every third person, where one person is in and then skip two; then the fourth person is in and skip two. Or, you could go every fourth person.

Often people prefer systematic samples to simple random samples because systematic samples are so much easier to take. It's easier than getting a whole list of people and assigning everyone a number or putting all the people's names in a hat. It's easier to take every fifth person or whatever you decide "k" should be.



HINT

The nice thing about a systematic sample is that it can be tailored to fit your sample size. If you wanted a sample of 25 from 500 individuals, you could sample every 20th person since 500 divided by 25 equals 20. So you would obtain your sample of 25 by sampling every 20th person.

IN CONTEXT

Suppose that you have 20 students in a class, and they're in rows, assigned to their desks randomly. If that were the case, you could count off every fourth student and have five students go up to the chalkboard to do a homework problem on the chalkboard.

1	2	3	✗	5
6	7	✗	9	10
11	✗	13	14	15
✗	17	18	19	✗

So, person one, two, and three don't have to do it. Person number four heads up to the chalkboard to work on a problem. Five, six, and seven don't have to do it, but number eight does. You can see the checkmarks to indicate the pattern and who needs to go up to the chalkboard.

What if they were alphabetized instead of randomly assigned?

Abbott	Acosta	Adams	Adamson ✗	Adler
Anderson	Bueller	Frye ✗	Grey	Jones
McClurg	Morris ✗	Peterson	Pickett	Rooney
Ruck ✗	Sara	Sheen	Stein	Ward ✗

By selecting say, Adamson, you automatically know who all the rest of the people are going to be. Since Adler is right next to Adamson, you know that Adler won't get chosen. Nor will Anderson or Bueller, but Frye will.

If these students were randomly assigned to the seats, picking Adamson would not predetermine who all the other people were going to be selected for the sample, but having them alphabetized impacts the random selection process.



TERM TO KNOW

Systematic Random Sample

A sampling method where every "k"th individual is selected for the sample (e.g. every 2nd, 4th, 20th individual).



SUMMARY

A simple random sample is the ideal sampling method if your goal is to obtain a representative sample. Sometimes, with big populations, it's not feasible to assign everyone a number or put everything into a hat so other sampling methods may be used. The random number generator is typically used with a calculator and is a fast way to calculate random "integers" without needing to assign same-number digits to each individual. The random number table is a more time-consuming method and generally used when technology is not available. A systematic sample can be similarly valid, and it is much easier to perform. It involves taking every "k"-th individual--however, the population must be randomly sorted before the systematic selection. Otherwise, it won't be considered random.

Good luck!



TERMS TO KNOW

Random Number Generator

A method of collecting a sample that utilizes technology to select random numbers corresponding to individuals in the population

Random Number Table

A method of collecting a sample to select random numbers corresponding to individuals in the population. Each individual is assigned a number, which are then selected from the table.

Simple Random Sample

A method of selection that guarantees that every sample of a certain size has an equal chance of being the selected sample

Systematic Random Sample

A sampling method where every "k"th individual is selected for the sample (e.g. every 2nd, 4th, 20th individual)