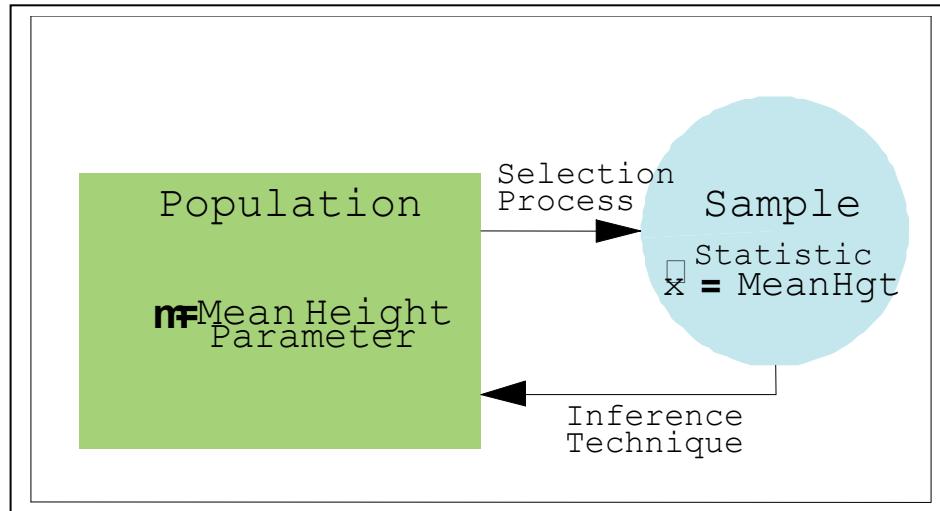


6:2a Introduction to Sampling



One of the basic ideas behind most statistics is the collection of information from a **simple random sample (SRS)** and the inference from that sample to a parameter in the population.

The purpose of a scientific approach to creating this simple random sample (SRS) is twofold;

- to eliminate bias
- to enable the researcher to quantify the distribution of the variance from sample to sample.

6:2b Introduction to Sampling

Examples of unsuspected bias are illustrated by the following scenarios.

Suppose that each member of this class was asked to pick a number from one to ten, based on probability each number should be selected approximately the same number of times and even numbers should be selected as often as odd numbers.

Practice, however, indicates that many groups of people will choose 7 more than predicted and choose 8 less often than predicted.

Suppose a deejay at a local radio station wanted to know the favorite song of his/her listening viewers and asked them to call in and vote for their favorite song. This type of sampling tends to be biased in favor of those with very strong opinions, frequently negative.

In 1990, the US government conducted a **census** designed to count the number of people living in the US. After the numbers were tabulated it was argued that the number of homeless people living in large cities was systematically **under reported**.

6:2c Introduction to Sampling

The first and second examples are considered **selection** biases. The third example illustrates undercoverage.

Another type of bias occurs when there is a **hidden structure** within the sampling frame used to generate the sample.

To eliminate these potential biases and to provide a basis for quantifying the distribution of the variance between the samples we need a method for generating a simple random sample.



6:2d Introduction to Sampling

A sample of size ' n ' is a '**simple random sample**' if and only if;

Each potential sample of size n from the population has an equal probability of being selected.

The selection or non-selection of an individual has no effect on the selection or non-selection of any other individual (they are independent).

Steps to Create a Simple Random Sample

- Decide on the population you wish to study
- Acquire a complete **sampling frame** for that population. A sampling frame is a list of all elements within the targeted population. **This step is not always possible.**

Before selecting a simple random sample you must assign labels to each and every data element in your sampling frame.

This is accomplished by numbering each element in order with whole numbers.

6:2e Introduction to Sampling

- 0 Frank Jones
- 1 Andre White
- 2 Mary Wilson
- 3 Omar Qatar
- 4 Jose Ramirez
- 5 Natasha Putin
- 6 Chantel Watkins
- 7 Michael Cullen
- 8 Melissa Norman
- 9 Alisha Hernandez

The next step is to determine the size of the required sample. This size is based on **the practical considerations of time, money and resources** and also on the theoretical requirements of the precision required in the end result. Generally there is **a trade off** between the desire on the part of the end user to be as precise as possible and the expenses required to achieve that precision.

6:2f Introduction to Sampling

Use either a table of random numbers or a random number generator to generate a list of 'n' random numbers. Usually by selecting more than the required number and eliminating duplicates.

Whichever method you use must be impersonal and impartial.

6:2g Introduction to Sampling

A table of random numbers.

Table of Random Digits

Line				
1	33586	72957	25989	14476
2	24956	38577	67318	84297

Decide on a starting spot.

Determine the number of digits you need, for example if there are 100 elements in your sampling frame you would need 2 digits 00 through 99, 1000 elements would require 3 digits, 000 through 999 etc.

Examine the table to find a label for the individual in your sampling frame.

For example, if there were 100 elements in the sampling frame you would be using 2 digits and the first five units selected would be 33, 58, 67, 29, 57. If you come to a two digit combination that has already been selected simply skip it.

6:2h Introduction to Sampling

If you come to a two digit combination that is larger than you need simply skip it.

Continue in this manner until you have generated the desired number of random numbers

6:2i Introduction to Sampling

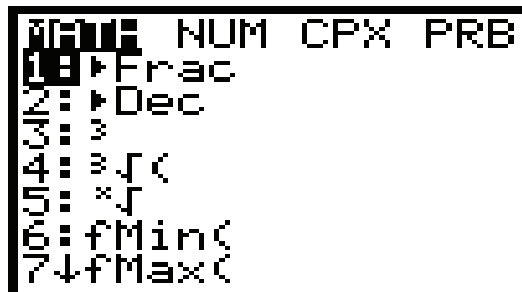
TI-83

A technique that has emerged over the past few years is to use the internal random number generator within calculators.

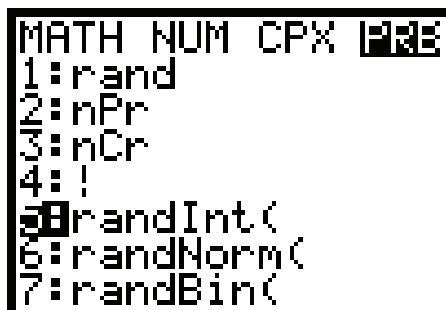
Place the calculator at the home screen

(<2nd> <MODE>)

Select MATH Menu



Select PRB Menu



6:2j Introduction to Sampling

Select randInt (*[option # 5]*

Enter lower bound, upper bound, numbers of entries then close the parenthesis and press the enter key

```
randInt(0,9,5)
{9 9 1 5 4}
```

This entry will generate 5 numbers between 0 and 9 inclusive.

Always select more entries than actually needed to provide for the inevitable duplication of random numbers as the above example indicates.

If you wish to insert these numbers into one of the STAT Lists (ex. List L1) insert the

<STO><2nd><1> keys before you press the enter key in the above step.

6:2k Introduction to Sampling

Excel -the third method involves using the internal capabilities of a spreadsheet program such as Excel. For this we use the function;

Rand ()

This function returns an evenly distributed random real number greater than or equal to 0 and less than 1. A new random number is returned every time the worksheet is calculated. This can be prevented by entering the 'F9' key after the cell entry

Remarks: To generate a random integer number between a and b, use: $=\text{INT}(\text{RAND}()*(b-a)+a)$