**Statistics for Beginners**

**Population Variables vs. Sample Variables**

For everyday use, we use the word “population” to refer to a group of people; for example, we often talk about the population of a country or a province. But in Stats, a **population** is the set of all the individuals of interest in a particular study. It can be anything in an investigation, for instance, cats, roses, ice cream and so on. Additionally, population can vary in size from extremely large to very small. It depends on how the investigator defines the population. For example, it can be the entire set (a well-defined collection of objects, i.e. a group of samples which have the same characteristics that are defined by the researcher) of the babies under two years old in the world or it can be the set of students at Nipissing. Even though a research aims at the entire population, it usually is impossible or impractical to collect all data of the population due to budgets, time, and resources ... Therefore, a set of individuals (**samples**) are selected from the population and intended to represent the population in a research study. For example, we are interested in the average height of women in Canada. Here, the population is all the women in Canada. However, it is impossible or impractical to obtain every woman’s height in Canada. What we can do, though, is select some women (for instance, one thousand women) as samples and use the data of the samples to estimate the average height of women in Canada.

|  |  |
| --- | --- |
| http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.pnghttp://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png

|  |
| --- |
| http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.pnghttp://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png…A thousand women selected from the the women in Canada (sample) |

 http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png http://static.guim.co.uk/sys-images/discussion/avatars/2010/08/22/JenBC/3b43a3b4-761a-4b3e-ac42-e5851cc90b06/60x60.png … |

All women in Canada (population)

Population and sample are distinct. How can we tell if a variable is a population variable or a sample variable in a question then? First, notations will tell you. We know that the symbol µ (population mean), σ (population standard deviation) and σ² (population variance) are used for population; and M (sample mean), S (sample standard deviation) and S² (sample variance) are used to indicate sample? sample. Secondly, population variables or sample variables sometimes are indicated in questions. Here is an example (a question in MATH 1257) as follows:

*A biologist knows that average length of a leaf of a certain full-grown plant is 4 inches. The* ***standard deviation of the population*** *is* ***0.6 inches****. A sample of 20 leaves of the type of plant given a new type of plant food had an average length of 4.2 inches. Is there reason to believe that the new food is responsible for a change in the grown of the leaves? Use α=0.01.*

In this question, the group of the all leaves of a certain full-grown plant is what the investigator wants to study so the group of all leaves of a certain full-grown plant is the population. The population mean is four inches which is the average length of a leaf of a certain full-grown plant. Obviously, it tells us that the 0.6 inches is the **population** standard deviation. And then the question says” a **sample** of 20 leaves had an average length of 4.2 inches”, so 20 inches is the sample mean.

Also, in the question, we notice that up to ***“0.6 inches”***, the question hasn’t mentioned anything about samples yet, so both “4 inches” and “0.6 inches” are population variables. Usually, if they are sample variables, the question will describe how many samples are selected from the population first and then tell you that the standard deviation and the mean of the samples.

Another example is as follows: 

*The true value of one type of degree or diploma cannot be quantitatively measured, but we can measure its relative impact on starting salary. Graduates from Quebec universities with a B.A. or B.Sc. degree have a mean annual starting salary of $28,300.* ***Sixty-five Quebec graduates*** *with a civil engineering degree are randomly selected. Their starting salaries have a mean of $36,300. If the standard deviation is $1670, use a 0.01 level of significance to test the claim that Quebec graduates with a civil engineering degree have a mean starting salary that is greater than the mean for graduates with a B.A. or B.Sc. degree from Quebec.*

This question talks all about the population until it describes that ***“Sixty-five Quebec graduates*** … are randomly selected.” Then it indicates that the mean of the sixty-five samples is $36,300 and the standard deviation of the sixty-five samples is $1670.

After you are able to tell the difference between population variables and sample variables, you can list the given variables in the question. The variable list will help you to decide which formula to apply and it will make it easier when you plug numbers into the formula. For example, I would list the known variables for the question above as follows:

µ=28.300

n=65

M=36,300

S=1670

Α=0.01

Reference

“*Statistics for the Behavioral Sciences*” by Frederick J Gravetter Larry B. Wallnau