



# REPRESENTATIVE SAMPLE – A CONCEPT TESTED IN GMAT CR

An e-GMAT Primer



**August 29, 2013**  
**Authored by: Chiranjeev Singh**

---

## INTRODUCTION

---

The objective of this article is to introduce a concept of statistics, which has been tested a number of times in the GMAT Critical Reasoning section and to use this concept to solve official CR arguments. Even though the concept is fairly simple to understand and apply, its ignorance leads to students' inability to select the correct option statements in such questions and ultimately to students paying the price in the form of an incorrect response.

---

## CR ARGUMENTS BASED ON RESEARCH FINDINGS

---

The concept that we will discuss in this article is: Representative sample. The understanding of this concept is tested in GMAT CR arguments, which draw conclusions based on the findings of research studies. An example of a research study could be a study which seeks to estimate the average score expectation of GMAT Test takers, just before they enter the exam hall and their actual average scores. Basically, a research study collects contextual data from the real world and draws conclusions on the basis of its analysis of the data.

In the context of GMAT, we need to understand one important characteristic of all research studies. The characteristic is that a research study deals with a sample but makes estimations about the entire population from which the sample is taken. To understand this characteristic, let's first understand the terms: sample and population.

---

## SAMPLE AND POPULATION

---

Let's understand these terms with an example of the research we talked about: a research which seeks to estimate the average score expectation of GMAT Test takers just before they enter their exam hall and their actual average scores.

In this research, the population is the set of all GMAT test takers. This is so because the study is interested in finding information about the set of all GMAT test takers.

Purpose = find the average GMAT SCORE of a population of GMAT test takers.

**Population:** Population is the entire group of entities (people, things, animals etc) we are interested in. It is the entire group we wish to understand or draw conclusions about.

However, we can see that collecting the required data for all the GMAT test takers may not be feasible for a research study because of the time and money required to do so. Therefore, generally a research study will select a small proportion of the GMAT test takers and collect data on these test takers.

The set of the test takers selected for the research is called 'sample'.

**Sample:** A sample is a group of units selected from a larger group (the population) or in other words, a sample is a subset of the population. A sample is generally selected for study because the population is too large to study in its entirety.

In the given research study, the sample could consist of

- 100 randomly selected people from fifteen different countries or
- all the GMAT test takers who take the GMAT at a particular test center over a one month period or
- all the GMAT test takers who are registered at GMATClub

So, a sample could be taken in several ways and as we can understand, the way a sample is taken will have an impact on the data we collect.

After the data is collected, based on this data (average expected score & actual average score), the research will make such estimates about the population i.e. all GMAT test takers.

---

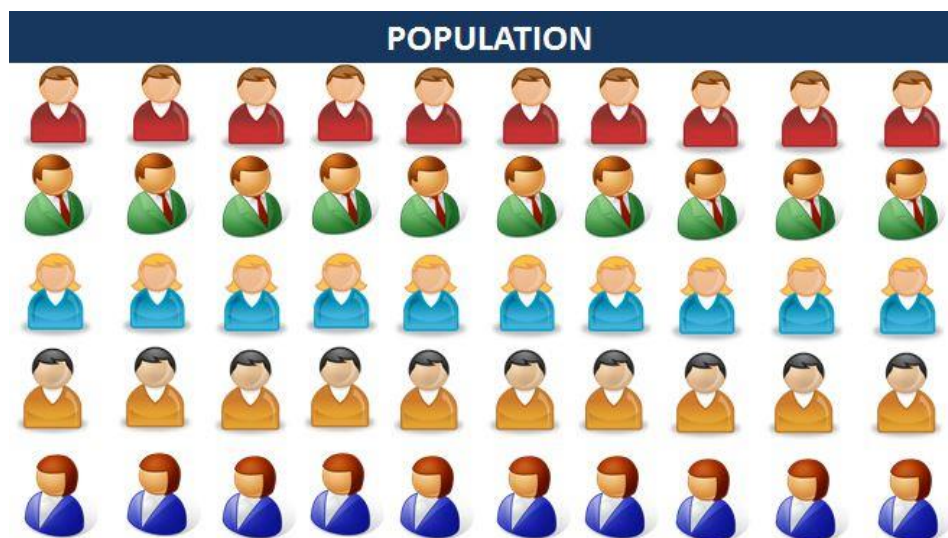
## MORE EXAMPLES OF SAMPLE AND POPULATION

---

---

### EXAMPLE 1

---



Let's consider a research study:

A study wants find the average income of a population of a city consisting of different income groups by measuring the average income of a section of the population.

What is the population in this case?

The population is the entire population of the city.

What is the sample?

The sample is the section of people whose average of income will be taken in the study.

## EXAMPLE 2

Let's consider another example of a research study:

A study wants to estimate the average salinity of Red sea by measuring the salinity of one million liters of water from the Red sea.

What is the population in this case?

The population will be the entire sea water of Red Sea.

What is the sample?

The sample is the one million liters of water that will be taken from Red sea.

## REPRESENTATIVE SAMPLE

As we have learnt so far, research studies use sample data to make estimates about the population data. Now, for sample data to give correct estimates of the population data, the chosen sample should be representative of the actual population or in other words, the chosen sample should give a true picture of the population.

A REPRESENTATIVE SAMPLE IS A SAMPLE WHICH GIVES A TRUE PICTURE OF THE POPULATION OR ONE WHOSE CHARACTERISTICS ARE IN LINE WITH THE CHARACTERISTICS OF THE POPULATION.

## EXAMPLE 1



If we want to find the average income of a population consisting of different income groups, then the sample must have people from all income groups in the same proportion as they constitute the population. So, if there are five sections in the population, each representing 20% of the population, then the sample must also have each of these sections constituting 20% of the sample. If the sample contains members of one group more than the members of the other group, then the average income of the sample will be quite different from the average income of the population.

For a sample to give a correct estimate of the population, the sample's characteristics have to be in line with the population's characteristics or in other words, the sample has to be representative of the actual population.

## EXAMPLE 2

---

If we seek to find the density of the sea water and we collect a sample from an area where the water density is abnormally high, would the sample data collected in this case give us the right estimates about the population data (i.e. sea water density)?

The answer is No. For a sample to give a correct estimate of the population, the sample's characteristics have to be in line with the population's characteristics or in other words, the sample has to be representative of the actual population.

In our sea water example, one way to make sure that the sample water is representative is that we collect water from different areas of the sea and mix them together to form a sample. In this way, we can make sure that we are not ending up selecting water which has abnormally higher or lower density than the sea water.

## EXAMPLE 3

---

Similarly, in the GMAT test takers research, if we somehow choose a sample of only high scorers, then the sample data collected will not give us the correct estimates of the population data. In such case, the average actual score and most probably, average expected too, would be higher than such figures for the population of all GMAT test takers.

## BUILDING STRENGTHENERS, WEAKENERS AND ASSUMPTIONS

---

We have learnt so far that for a sample to give correct estimates about the population, the sample has to be representative of the population.

### WEAKENER

---

Now, suppose we have an argument which draws a conclusion based on the findings of a research study. In such case, what would happen if we say that the sample used by the research study was not representative of the actual population?

The answer is that our trust in the conclusion will be significantly weakened since the conclusion depended on the validity of the research findings. Therefore, a statement suggesting that the sample used in conducting the research was unrepresentative would be a valid weakener for the argument.

### STRENGTHENER

---

Now, suppose if we say just the opposite - we say that the sample used by the research is actually representative of the population - would our belief in the conclusion be strengthened?

The answer is Yes. Now, with this additional information, we are surer of the research findings and hence our belief in conclusion which depends on the findings has increased.

---

### ASSUMPTION

---

Based on the weakener we have discovered, can you think of an assumption made in the argument which draws a conclusion based on the research findings?

The assumption is that the sample used in the research was representative of the actual population. This assumption is **required** because if this is not true, the research finding will not be believable and the conclusion will break down.

Now, let's look at two OG questions which use this understanding of representative sampling: one weaken question and one assumption question.

---

### OG QUESTION – WEAKEN

---

Solve this question yourself before reading the analysis:

A study of high blood pressure treatments found that certain meditation techniques and the most commonly prescribed drugs are equally effective if the selected treatment is followed as directed over the long term. Half the patients given drugs soon stop taking them regularly, whereas eighty percent of the study's participants who were taught meditation techniques were still regularly using them five years later. Therefore, the meditation treatment is the one likely to produce the best results.

Which of the following, if true, most seriously weakens the argument?

- A. People who have high blood pressure are usually advised by their physicians to make changes in diet that have been found in many cases to reduce the severity of the condition.
- B. The participants in the study were selected in part on the basis of their willingness to use meditation techniques.
- C. Meditation techniques can reduce the blood pressure of people who do not suffer from high blood pressure.
- D. Some of the participants in the study whose high blood pressure was controlled through meditation techniques were physicians.
- E. Many people with dangerously high blood pressure are unaware of their condition.

---

### ANALYSIS

---

The answer to the question is option B. Let's understand this:

What is the population for the study mentioned?

The population is the set of all high BP (Blood Pressure) patients.

Sample is the set of all the participants of the study.

Now, if we say that the people chosen in the sample were those who were more willing to use meditation techniques, can we call this sample a representative sample?

The answer is No. The sample is rather biased i.e. it has people who were more willing to use meditation techniques than the prescribed drugs. A representative sample would have been the one that was selected without considering the willingness of the people of one treatment over the other.

Now, since the sample used in the study is not representative, we cannot believe in the results of study. Since we cannot believe in the results of the study, there is no basis to believe in the conclusion of the argument, which used the research findings as its premises. Rather, in the study conducted using the biased sample, it was expected that people in the study would use meditation techniques over longer term than medications because the study selected only those people who were more willing to use meditation techniques.

So, even if in general population, people might not actually use meditation techniques over the longer term than medications, the study will still support meditation techniques because the chosen sample was biased to favor meditation techniques.

Therefore, we can see that option B creates doubts on the conclusion and hence, is a valid weakener.

---

## OG QUESTION – ASSUMPTION

---

The Earth's rivers constantly carry dissolved salts into its oceans. Clearly, therefore, by taking the resulting increase in salt levels in the oceans over the past hundred years and then determining how many centuries of such increases it would have taken the oceans to reach current salt levels from a hypothetical initial salt-free state, the maximum age of the Earth's oceans can be accurately estimated.

Which of the following is an assumption on which the argument depends?

- A. The quantities of dissolved salts deposited by rivers in the Earth's oceans have not been unusually large during the past hundred years.
- B. At any given time, all the Earth's rivers have about the same salt levels.
- C. There are salts that leach into the Earth's oceans directly from the ocean floor.
- D. There is no method superior to that based on salt levels for estimating the maximum age of the Earth's oceans.
- E. None of the salts carried into the Earth's oceans by rivers are used up by biological activity in the oceans

---

## ANALYSIS

---

The answer to the question is option A. Let's understand this:

This question is quite tricky since we are making two levels of estimations.



One estimate is what we are given in the passage i.e. estimate of the earth's age.

However, to get an estimate of the earth's age, we need to get an estimate of the rate of increase of salt level. This is because as given in the passage, we are going to calculate the maximum age of the earth using the formula:

Current salt level of the ocean/rate of increase of salt level.

So, suppose if we have the current salt level as 100 and we estimate the rate of increase of salt level as 2, then we can say that the maximum age of the earth is  $100/2$  i.e. 50.

Therefore, to estimate the maximum age of the earth, we need to estimate the rate of increase of salt level. To get a correct estimate of the earth's age, we need to have the average rate of increase of salt level from the beginning of the formation of the earth till now.

Since, we are looking at a period from the beginning of the earth till now, the population is this entire period. The sample in this case is the last 100 years.

From our understanding of representative samples, we know that to get any correct estimates, the rate of increase of salt level in the last 100 years should be representative of the rate of increase of salt level of the entire period.

Option A communicates the same message by saying that the sample is not unrepresentative i.e. increase in salt levels over the past 100 years have not been unusually large. Moreover, if we negate option A, we have that the quantities of dissolved salts deposited by rivers in the Earth's oceans have been unusually large during the past hundred years. This negated statement means that the sample is unrepresentative. Thus, negating option A breaks down the conclusion.

Therefore, option A is the required assumption.

Here, I just want to talk a bit about option E because that is the most confusing option statement. For option E to be an assumption, it must pass negation test. When we negate option E, we have that some of the salts carried into the Earth's oceans by rivers are used up by biological activity in the oceans. Does that break down the conclusion? No. If the rate of consumption of salt by the biological activity is not unusually higher or lower in the last 100 years than it has been from the beginning of the earth, we have a representative sample and thus, we'll have a correct estimate of the earth's age. This means that even if some of the salts are used by the biological activity, we can still accurately estimate the earth's age. Therefore, negation option E does not break down the conclusion.

---

## TAKE AWAYS

---

So, what have we learnt from this article:

1. For a sample to provide a correct estimate of the population, it must be a representative sample i.e. it must have the same characteristics as the population
2. An argument which draws a conclusion based on a research findings:
  - a. Can be weakened by suggesting that the sample chosen was not representative of the population



- b. Can be strengthened by suggesting that the sample chosen was indeed representative of the population
- c. Is based on the assumption that the sample chosen in the study was representative of the population.

## EXERCISE QUESTIONS

---

1. Often patients with ankle fractures that are stable, and thus do not require surgery, are given follow-up x-rays because their orthopedists are concerned about possibly having misjudged the stability of the fracture. When a number of follow-up x-rays were reviewed, however, all the fractures that had initially been judged stable were found to have healed correctly. Therefore, it is a waste of money to order follow-up x-rays of ankle fracture initially judged stable.

Which of the following, if true, most strengthens the argument?

- A. Doctors who are general practitioners rather than orthopedists are less likely than orthopedists to judge the stability of an ankle fracture correctly.
  - B. Many ankle injuries for which an initial x-ray is ordered are revealed by the x-ray not to involve any fracture of the ankle.
  - C. X-rays of patients of many different orthopedists working in several hospitals were reviewed.
  - D. The healing of ankle fractures that have been surgically repaired is always checked by means of a follow-up x-ray.
  - E. Orthopedists routinely order follow-up x-rays for fractures of bone other than ankle bones.
2. Frobisher, a sixteenth-century English explorer, had soil samples from Canada's Kodlunarn Island examined for gold content. Because high gold content was reported, Elizabeth I funded two mining expeditions. Neither expedition found any gold there. Modern analysis of the island's soil indicates a very low gold content. Thus the methods used to determine the gold content of Frobisher's samples must have been inaccurate.

Which of the following is an assumption on which the argument depends?

- A. The gold content of the soil on Kodlunarn Island is much lower today than it was in the sixteenth century.
- B. The two mining expeditions funded by Elizabeth I did not mine the same part of Kodlunarn Island.
- C. The methods used to assess gold content of the soil samples provided by Frobisher were different from those generally used in the sixteenth century.
- D. Frobisher did not have soil samples from any other Canadian island examined for gold content.
- E. Gold was not added to the soil samples collected by Frobisher before the samples were examined.