

## CMD 305 - COURSE

### (RESEARCH METHODOLOGY & BIOSTATISTICS)

#### Tutorial Topic: sampling techniques (with solutions)

---

Q1) What do you mean by 'sample' and population? Explain with a simple example.

- **Solution:** A small portion/group of subjects selected from a wider group of subjects is called a sample. This wider group is called population.

Example : Say, in a particular hospital 1000 deliveries occurred in a particular year and out of these we take 100 deliveries for our research study. These 1000 deliveries is our population and these 100 deliveries is our sample.

Q2. Why do you study only a sample of patients?

- **Solution:** To collect the data from all the subjects of the population, it takes long time and also we need lot of resources and trained persons to collect data, so, we study only a sample of patients.

Q3) a) What do you call that sample where subjects are selected as you wish?  
b) What do you call that sample where subjects are selected without any bias?

**Solution:** a) Convenient sample or non random sample.  
b) Random sample.

Q4) Give names of some of the random sampling techniques you know.

**Solution:** Simple random sampling, Systematic random sampling, Stratified random sampling, Cluster random sampling and multistage random sampling.

Q5) In a big hospital, every year 500 cases of MI (myocardial infarction) are reported. We want to study their physiological profile-their BP, cholesterol level, lipoprotein level, BMI, etc. Resources permit us to do investigations only for 50 cases. How do you select a **simple random sample(SRS)** of 50 cases out of these 500 cases. Explain the crude way as well as easy way to select this sample.

**Solution:** We will write ID numbers of these 500 cases in 500 similar looking slips and roll them and put in a bowl and shuffle well and then take 50 slips one by one. The patients whose ID numbers are picked up is our sample. This method of sampling is called simple random sampling. This is the crude way and difficult to do. Easy way is take 50 random numbers within 1 to 500 from the computer or from the random number tables and the patients whose ID numbers are selected, will be our sample.

Q6) A researcher wants to take a random sample of 100 cases from 1000 deliveries that occurred in maternity hospital in the last year. He has taken one random number out of 1 and 10 say, 5. Then he took a case having ID No.5. Then he took cases having ID numbers 15,25,35,45.....995 as his sample. What method of sampling the researcher adopted here?

**Solution :** here, the researcher has adopted systematic sampling.

Q7) Consider one more hospital where 1000 MI cases were reported last year. He wants to do a study on these cases. As these number of cases is large, he wants to take a sample of 100 cases. And also, as the physiological parameters of these cases would be different in overweight and less weight cases, the researcher wants to divide these 1000 MI cases into two groups one with overweight/obese(that is  $BMI \geq 25$ ) cases and the other less weight( $BMI < 25$ ) cases and that both these groups to represent in his sample of 100 cases. Then he took a sample of 50 patients at random from each of these two groups.

- a) What are these two groups called in sampling?
- b) What is the sampling method adopted here to select a sample of 100 cases? Why did the researcher adopt this sampling method?

**Solution :** a) These groups are called strata in sampling.  
b) This method of sampling is called stratified sampling. He adopted this sampling method because that both the strata that is, overweight and less weight cases to be represented in the sample.

Q8) It was decided to estimate prevalence of diabetes in KSA. He had limited resources. So, he divided entire KSA into 5 regions as north, south, east, west and central. Then he made 20 contiguous geographical areas (clusters) in each of these five regions. Then he selected two clusters at random from each of five regions. He found 5000 eligible individuals from these five selected areas. And he collected relevant data from all these individuals.

- a) What type of sampling method he used here?
- b) Why did he adopt this method?

**Solution:**

- a) Cluster sampling
- b) He used cluster sampling because he had limited resources and he does not need sampling frame that is list of all the individuals of entire KSA, which is difficult to get. It is enough if he has list of clusters, and he could collect the

data from all the individuals of the selected clusters only. In this way, he saves lot of resources by not travelling widely to take a simple random sample.

Q9) Health authorities asked an epidemiologist to find out the prevalence of anemia in high school children of standard VI to X in a district of an African country. There are 60 schools in this district. And each school has standard VI to X classes. He wanted to use multistage sampling method to estimate the prevalence of anemia in high school children of standard VI to X in that district? How he would have done multistage sampling method in this situation.

**Solution:** First, he needs only the list of these 60 schools. In the first stage he can select 5 schools among 60 schools at random, and from each of the selected school, out of five standards VI to X, select two standards at random. This is second stage of selection. And from each of the selected standard select 20 students at random. This the third stage of selection. So our sample consists of 200 students (5x2x20). This is his sample of students from whom he has to collect data to estimate the prevalence of anemia of high school children of that district.