RESEARCH METHODOLOGY

LECTURE-30

SOME FUNDAMENTAL DEFINITIONS

Before we talk about details and uses of sampling, it seems appropriate that we should be familiar with some fundamental definitions concerning sampling concepts and principles.

1. Universe/Population: From a statistical point of view, the term 'Universe'refers to the total of the items or units in any field of inquiry, whereas the term 'population' refers to the total of item about which information is desired. The attributes that are the object of study are referred to as characteristics and the units possessing them are called as elementary units The aggregate of such is generally described as population. Thus, all units in any field of inquiry constitute universe an all elementary units (on the basis of one characteristic or more) constitute population. Quit often, w do not find any difference between population and universe, and as such the two terms are taken a interchangeable. However, a researcher must necessarily define these terms precisely The population or universe can be finite or infinite.

2. Sampling frame: The elementary units or the group or cluster of such units may form the basis sampling process in which case they are called as sampling units. A list containing all such sampling units is known as sampling frame. Thus sampling frame consists of a list of items from which th sample is to be drawn. If the population is finite and the time frame is in the present or past, then it is possibe for the frame to be identical with the population. In most cases they are not identical because it is often impossible to draw a sample directly from population. As such this frame is either constructed by a researcher for the purpose of his study or may consist of some existing list of the population. For instance, one can use telephone directory as a frame for conducting opinion survey in a city. Whateve the frame may be, it should be a good representative of the population.

1. Sampling design: A sample design is a definite plan for obtaining a sample from the sampling frame. It refers to the technique or the procedure the researcher would adopt in selecting som sampling units from which inferences about the population is drawn. Sampling design is determine

before any data are collected.

4. Statisitc(s) and parameter(s): A statistic is a characteristic of a sample, whereas a parameter is a characteristic of a population. Thus, when we work out certain measures such as mean, median, mode or the like ones from samples, then they are called statistic(s) for they describe the characteristics of a sample. But when such measures describe the characteristics of a population, they are know as parameter(s). For instance, the population mean bmg is a parameter,whereas the sample mean (X) is a statistic. To obtain the estimate of a parameter from a statistic constitutes the prim objective of sampling analysis.

2. Sampling error: Sample surveys do imply the study of a small portion of the population and a such there would naturally be a certain amount of inaccuracy in the information collected. This inaccuracy may be termed as sampling error or error variance.

6. Precision: Precision is the range within which the population average (or other parameter) wil lie in accordance with the reliability specified in the confidence level as a percentage of the estimate \pm or as a numerical quantity. For instance, if the estimate is Rs 4000 and the precision desired is \pm 4%, then the true value will be no less than Rs 3840 and no more than Rs 4160. This is the range (Rs 3840 to Rs 4160) within which the true answer should lie. But if we desire that the estimate should not deviate from the actual value by more than Rs 200 in either direction, in that case the range would be Rs 3800 to Rs 4200.

7. Confidence level and significance level: The confidence level or reliability is the expected of times that the actual value will fall within the stated precision limits. Thus, if we take

a confidence level of 95%, then we mean that there are 95 chances in 100 (or .95 in 1) that th sample results represent the true condition of the population within a specified precision range against 5 chances in 100 (or .05 in 1) that it does not. Precision is the range within which the answer may vary and still be acceptable; confidence level indicates the likelihood that the answer will fall within that range, and the significance level indicates the likelihood that the answer will fall outside that range. We can always remember that if the confidence level is 95%, then the significance level will be (100 - 95) i.e., 5%; if the confidence level is 99%, the significance level is (100 - 99) i.e., 1%, and so on. We should also remember that the area of normal curve within precision limits for the specified confidence level constitute the acceptance region and the

area of the curve outside these limits in direction constitutes the rejection regions

8. Sampling distribution: We are often concerned with sampling distribution in sampling analysis If we take certain number of samples and for each sample compute various statistical measures as mean, standard deviation, etc., then we can find that each sample may give its own value for the statistic under consideration. All such values of a particular statistic, say mean, together with their relative frequencies will constitute the sampling distribution of the particular statistic, say mean.

Accordingly, we can have sampling distribution of mean, or the sampling distribution of standard deviation or the sampling distribution of any other statistical measure. It may be noted that each item in a sampling distribution is a particular statistic of a sample. The sampling distribution tends quite closer to the normal distribution if the number of samples is large. The significance of sampling distribution follows from the fact that the mean of a sampling distribution is the same as the mean of the universe. Thus, the mean of the sampling distribution can be taken as the mean of the universe.