# Lecture 8
# SAMPLING

By. Dr. Rafiq Mansoor

# Sampling

- The items so selected constitute what is technically called a sample, their selection process or technique is called sample design and the survey conducted on the basis of sample is described as sample survey.

# SOME FUNDAMENTAL DEFINITIONS

1. *Universe/Population: 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* of sampling process in which case they are called as sampling units.
3. *Sampling design: A sample design is a definite plan for obtaining a sample from the sampling* frame.
4. *Statisitc(s) and parameter(s): A statistic is a characteristic of a sample, whereas a parameter is* a characteristic of a population.

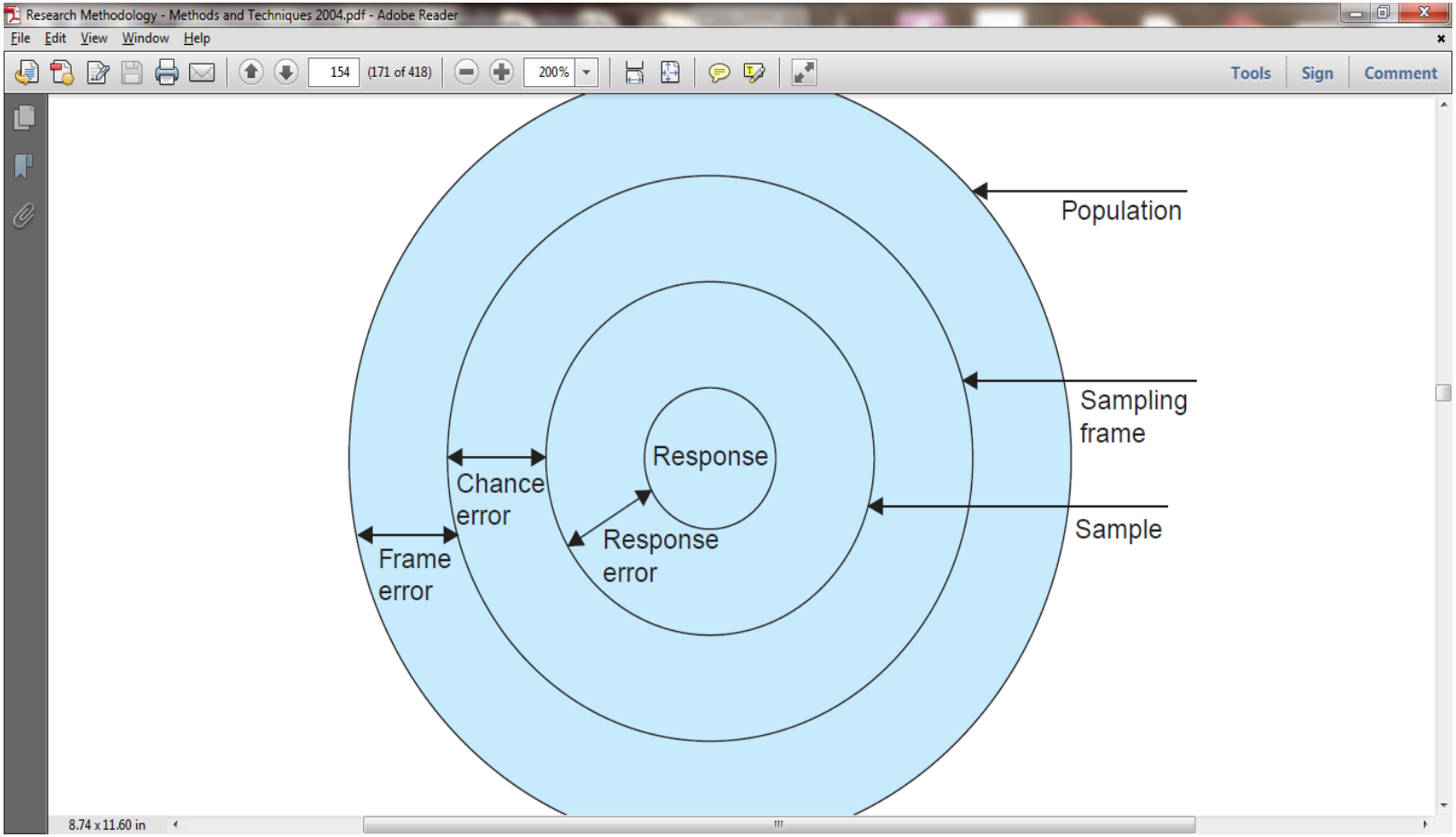
# SOME FUNDAMENTAL DEFINITIONS

5. *Sampling error: Sample surveys do imply the study of a small portion of the population and* as such there would naturally be a certain amount of inaccuracy in the information collected.

The meaning of sampling error can be easily understood from the following diagram:

# Fig. 8.1

## Sampling error = Frame error + Chance error + Response



# SOME FUNDAMENTAL DEFINITIONS

**6. Precision:**

- **Precision is the range within which the population average (or other parameter) will lie in accordance with the reliability specified in the confidence level as a percentage of the estimate ± or as a numerical quantity.**

- **For instance, if the estimate is Rs 4000 and the precision desired is ± 4%, then the true value will be no less than Rs 3840 and no more than Rs 4160.**

# SOME FUNDAMENTAL DEFINITIONS

***7. Confidence level and significance level:***

- **The confidence level or reliability is the expected percentage 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 the sample results represent the true condition of the population within a specified precision range against5 chances in 100 (or .05 in 1) that it does not.**

# SOME FUNDAMENTAL DEFINITIONS

**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 such as mean, standard deviation, etc.,* then we can find that each sample may give its own value for the statistic under consideration.**

# IMPORTANT SAMPLING DISTRIBUTIONS

- Some important sampling distributions, which are commonly used, are:
- (1) sampling distribution of mean;
- (2) sampling distribution of proportion;
- (3) student's *'t' distribution;*
- *(4) F distribution; and*
- (5) Chi-square distribution.

# Central Limit Theorem

- from a normal population, the means of samples drawn from such a population are themselves *normally distributed*.
- But when sampling is not from a normal population, the size of the sample plays a critical role. When *n is small, the shape of the distribution will depend largely on the* shape of the parent population, but as *n gets large (n > 30), the thape of the sampling distribution will* become more and more like a normal distribution, irrespective of the shape of the parent population.

# Central Limit Theorem

- **"The significance of the central limit theorem lies in the fact that it permits us to use sample statistics to make *inferences* about population parameters without knowing anything about the shape of the frequency distribution of that population other than what we can get from the sample."**

# SAMPLING THEORY

- **Sampling theory is a study of relationships existing between a population and samples drawn from the population.**

- **Sampling theory is designed to attain one or more of the following objectives:**
- **(*i) Statistical estimation:***

  **The estimate can either be a point estimate or it may be an interval estimate.**

# SAMPLING THEORY

- ***(ii) Testing of hypotheses:***

  ***The second objective of sampling theory is to enable us to decide* whether to accept or reject hypothesis;**

- ***(iii) Statistical inference:***

  ***Sampling theory helps in making generalization about the population/* universe from the studies based on samples drawn from it. It also helps in determining the accuracy of such generalizations.**

# CONCEPT OF STANDARD ERROR
# The standard deviation

- The standard deviation of sampling distribution of a statistic is known as its standard error (S.E) and is considered the key to sampling theory.
- The utility of the concept of standard error in statistical induction arises on account of the following reasons:

  1. The (S.E) helps in testing whether the difference between observed and expected frequencies could arise due to chance.
  The criterion usually adopted is that if a difference is less than *3 times* the S.E., the difference is supposed to exist as a matter of chance and if the difference is equal to or more than 3 times the S.E., chance fails to account for it, and we conclude the difference as significant difference. This criterion is based on the fact that at $X \pm 3$ (S.E.) the normal curve covers an area of 99.73 per cent.

- **2. The standard error gives an idea about the reliability and precision of a sample. The *smaller the S.E., the greater the uniformity* of sampling distribution and hence, greater is the reliability of sample.**

- **Conversely, the *greater the S.E., the greater* the difference between observed and expected frequencies. In such a situation the unreliability of the sample is greater.**

- **3. The standard error enables us to *specify the limits* within which the parameters of the population are expected to lie with a specified degree of confidence. Such an interval is usually known as *confidence interval*.**

# ESTIMATION

- In most statistical research studies, population parameters are usually unknown and have to be estimated from a *sample*.

# Sample size and its detemination

- **In sampling analysis the most ticklish *question* what should be the *size of the sample* or how large or small should be *'n'? If the sample size ('n') is too small, it may not serve to achieve the objectives* and if it is too large, we may incur huge cost and waste resources.**

# DETERMINATION OF SAMPLE SIZE THROUGH THE APPROACH BASED ON PRECISION RATE & CONFIDENCE LEVEL

- To begin with, it can be stated that whenever a sample study is made, there arises some sampling error which can be controlled by selecting a sample of adequate size.