

## Question Paper: Sampling Theory

### Instructions:

- **Time Allowed:** 2 Hours
- **Total Marks:** 60
- **Answer all questions.**
- **Marks are indicated next to each question.**

### Section A: Objective Type Questions (20 Marks)

**Instructions:** Choose the best answer for each multiple-choice question. Each question is worth 1 mark.

1. The sampling error decreases with:
  - a) Decrease in sample size
  - b) Increase in sample size
  - c) Homogeneity of the population
  - d) All of the above**Answer:** (b)
2. Which of the following is *not* a type of probability sampling?
  - a) Simple Random Sampling
  - b) Stratified Random Sampling
  - c) Convenience Sampling
  - d) Cluster Sampling**Answer:** (c)
3. In stratified sampling, the strata should be:
  - a) Homogeneous within and heterogeneous between
  - b) Heterogeneous within and homogeneous between
  - c) Homogeneous both within and between

d) Heterogeneous both within and between

**Answer:** (a)

4. The standard deviation of the sampling distribution of a statistic is called:

- a) Standard Error
- b) Margin of Error
- c) Bias
- d) Confidence Interval

**Answer:** (a)

5. A sampling frame is:

- a) A list of all the sampling methods.
- b) A summary of the various stages involved in designing a survey.
- c) A list of all the units in the population from which a sample will be selected.
- d) A framework for data analysis.

**Answer:** (c)

6. Systematic sampling is appropriate when:

- a) The population is homogeneous.
- b) The population exhibits a trend.
- c) The population has distinct groups.
- d) When data is already organized in a sequence.

**Answer:** (d)

7. Which of the following is true about the cluster sampling?

- a) It is cost-effective.
- b) It requires a complete list of population members.
- c) It is more statistically precise than simple random sampling.
- d) It is always unbiased.

**Answer:** (a)

8. Quota sampling is an example of:

- a) Probability sampling
- b) Non-probability sampling
- c) Stratified sampling
- d) Cluster sampling

**Answer:** (b)

9. If we increase the confidence level, the width of the confidence interval:

- a) Decreases
- b) Increases
- c) Remains the same
- d) Depends on the sample size

**Answer:** (b)

10. The central limit theorem states that the sampling distribution of the sample mean is approximately normal when:

- a) The sample size is large enough.
- b) The population is normally distributed.

- c) The population variance is known.
- d) The population size is small.

**Answer:** (a)

11. What is the purpose of sampling?

- a) To study the entire population
- b) To gather data from a representative subset of the population
- c) To eliminate bias in research
- d) To create a perfect representation of the population

**Answer:** (b)

12. What is the potential problem with volunteer sampling?

- a) It is too expensive.
- b) It is always perfectly representative.
- c) It is prone to bias.
- d) It is too time-consuming.

**Answer:** (c)

13. Which type of sampling is particularly useful when the population is geographically dispersed?

- a) Simple Random Sampling
- b) Stratified Sampling
- c) Cluster Sampling
- d) Systematic Sampling

**Answer:** (c)

14. What does 'n' typically represent in sampling theory?

- a) Population size
- b) Sample size
- c) Standard deviation
- d) Mean

**Answer:** (b)

15. What happens to the standard error as the sample size increases?

- a) It increases.
- b) It decreases.
- c) It stays the same.
- d) It fluctuates randomly.

**Answer:** (b)

16. Which of the following is the main advantage of using a larger sample size?

- a) Reduced cost
- b) Increased accuracy
- c) Faster data collection
- d) Easier analysis

**Answer:** (b)

17. Snowball sampling is most commonly used in:

- a) Large-scale surveys

- b) Studies of hidden populations
- c) Clinical trials
- d) Market research

**Answer:** (b)

18. In systematic sampling, what is the 'sampling interval'?

- a) The number of strata
- b) The number of clusters
- c) The fixed interval between selected elements
- d) The margin of error

**Answer:** (c)

19. What type of error cannot be eliminated by increasing the sample size?

- a) Sampling error
- b) Non-sampling error
- c) Standard error
- d) Type I error

**Answer:** (b)

20. Which of the following distributions is frequently used as an approximation for the sampling distribution of the sample mean when the population standard deviation is unknown and the sample size is small?

- a) Normal distribution
- b) t-distribution
- c) Chi-square distribution
- d) F-distribution

**Answer:** (b)

## Section B: Subjective Type Questions (40 Marks)

**Instructions:** Answer the following questions in detail.

1. **(8 Marks)**

- a) Define 'sampling' and explain why sampling is necessary in statistical inference. (4 Marks)
- b) Differentiate between a 'population' and a 'sample'. Give examples. (4 Marks)

**Answer:**

- **(a)** Sampling is the process of selecting a subset of individuals from a larger population to estimate characteristics of the whole population. Sampling is necessary because it is often impractical, costly, or impossible to study the entire population. It allows us to make inferences about the population based on the characteristics of the sample.
- **(b)** A **population** is the entire group of individuals, objects, or events that are of interest in a study. For example, all registered voters in a country, all students in a university, or all cars produced by a factory in a year. A **sample** is a subset of the population that is selected for study. For example, a random selection of 1000

registered voters from a country, 200 students from a university, or 50 cars from the factory's production line.

2. **(8 Marks)**

- a) Explain the difference between probability and non-probability sampling. (4 Marks)
- b) Describe two different probability sampling methods, highlighting their advantages and disadvantages. (4 Marks)

**Answer:**

- **(a) Probability sampling** methods involve random selection, ensuring that each member of the population has a known (non-zero) chance of being selected into the sample. This allows for statistical inferences about the population. **Non-probability sampling** methods do not use random selection. The probability of selecting a particular element is unknown and may be influenced by the researcher's subjective judgment. These methods are often used for exploratory research but do not allow for reliable statistical generalizations.
- **(b)**
  - **Simple Random Sampling (SRS):**
    - **Advantage:** Easiest method, free from classification error, requires minimum knowledge of population.
    - **Disadvantage:** Not suitable for widely spread items, requires complete updated list of population, may result in a sample that does not accurately represent the population if the sample size is small.
  - **Stratified Random Sampling:**
    - **Advantage:** Ensures representation of all subgroups in the population, can reduce sampling error compared to SRS.
    - **Disadvantage:** Requires knowledge of the population structure, can be more complex to implement than SRS.

3. **(8 Marks)**

- a) Define 'sampling error' and explain its causes. (4 Marks)
- b) Explain what is meant by the 'Central Limit Theorem' and its importance in sampling theory. (4 Marks)

**Answer:**

- **(a) Sampling error** is the difference between a sample statistic (e.g., sample mean) and the corresponding population parameter (e.g., population mean). It occurs because a sample is only a subset of the population and may not perfectly reflect the population's characteristics. Causes of sampling error include:
  - Chance variation (randomness in the selection process).
  - Sample size (smaller samples tend to have larger sampling errors).
  - Population variability (more heterogeneous populations tend to result in larger sampling errors).

- **(b) The Central Limit Theorem (CLT)** states that, regardless of the shape of the population distribution, the sampling distribution of the sample mean will approach a normal distribution as the sample size increases. The importance of the CLT is that it allows us to use normal distribution theory to make inferences about the population mean, even if the population is not normally distributed, provided the sample size is sufficiently large (generally,  $n \geq 30$  is considered sufficient).

4. **(8 Marks)**

a) Explain the concept of standard error. How is it related to sample size and population variability? (4 Marks)

b) Discuss the factors that should be considered while determining the sample size for a research study. (4 Marks)

**Answer:**

- **(a) Standard error** is the standard deviation of the sampling distribution of a statistic. It measures the variability or precision of the sample statistic as an estimator of the population parameter. Standard error is **inversely proportional** to the square root of the sample size ( $n$ ). As the sample size increases, the standard error decreases, indicating greater precision. Standard error is **directly proportional** to the population variability (standard deviation,  $\sigma$ ). More variable populations result in larger standard errors.
- **(b) Factors to consider when determining sample size:**
  - **Population variability:** Higher variability requires larger sample sizes.
  - **Desired precision (margin of error):** Smaller margins of error require larger sample sizes.
  - **Confidence level:** Higher confidence levels require larger sample sizes.
  - **Budget:** Budget constraints may limit the feasible sample size.
  - **Expected effect size:** Smaller effect sizes may require larger sample sizes.
  - **Statistical power:** Desired power to detect a statistically significant effect, larger power needs larger sample size.
  - **Type of analysis:** More complex statistical analyses may require larger sample sizes.

5. **(8 Marks)**

a) What is cluster sampling and when is it most appropriate to use? Provide an example. (4 Marks)

b) Explain the difference between stratified sampling and quota sampling. (4 Marks)

**Answer:**

- **(a) Cluster sampling** involves dividing the population into groups (clusters), randomly selecting some of these clusters, and then sampling all or a subset of individuals within the selected clusters. It is most appropriate when:
  - The population is naturally divided into clusters.
  - It is costly or difficult to create a complete list of individuals in the population.

- Geographical dispersion of the population is high.
- **Example:** A researcher wants to study the academic performance of high school students in a large city. Instead of randomly selecting students from all schools in the city, they could randomly select a few schools (clusters) and then survey all students in those selected schools.
- **(b)**
  - **Stratified sampling** is a probability sampling method where the population is divided into subgroups (strata) based on shared characteristics, and then a random sample is taken from each stratum. The number of samples selected from each stratum are usually in proportion to the size of the stratum within the population.
  - **Quota sampling** is a non-probability sampling method where the sample is selected based on pre-determined quotas for different subgroups. The quotas are designed to ensure that the sample reflects the population's characteristics in terms of certain variables (e.g., age, gender, ethnicity). However, selection within each quota is not random and may be based on convenience or judgment. **Key difference:** Stratified sampling uses random selection *within* each stratum, while quota sampling does not guarantee random selection.
  - In stratified sampling, the selection is random after forming the subgroups. Quota sampling fills quotas based on researcher's judgment.