DEPARTMENT OF INFORMATION & COMMUNICATION TECHNOLOGY
EXERCISE | 2 2024/2025

COURSE CODE: DFC30323	**COURSE NAME: STATISTICS AND PROBABILITY**	
EXERCISE: SAMPLING AND ESTIMATION (CLO1, C2)		
NAME:		
REGISTRATION NO:	**CLASS:**	**DURATION: 2 HOURS**

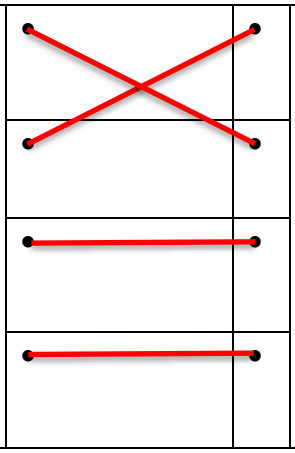
QUESTION

1. Define sampling and explain why it is used in statistics. Sampling is the process of selecting a subset of individuals from a larger population to study. It is used in statistics to make inferences about a population without studying every individual, saving time and resources.	2 Marks
2. Classify each sample as simple random, systematic, stratified, or cluster. a) Out of every 50 cars manufactured, is checked to determine its gear. Systematic Sampling – Every 50th car is checked, meaning there is a fixed interval between selections. b) Out of every 10 customers entering a shopping mall, is asked to select his favourite store. Systematic Sampling – Every 10th customer is selected, following a predetermined pattern. c) Assistant professors are selected using random numbers to determine annual salaries. Simple Random Sampling – Assistant professors are chosen using random numbers, ensuring each has an equal chance of selection.	3 Marks
3. Identify the population and sample in the following scenarios: a) A researcher surveys 500 students from a university with 20,000 students about their study habits. Sample: 500 students surveyed. Population: All 20,000 students in the university. b) A food company selects 100 bottles of juice from a production batch of 10,000 bottles to test quality. Sample: 100 bottles of juice. Population: All 10,000 bottles in the batch. c) A car company checks every 20th car produced in a day to inspect for defects. Sample: The inspected cars. Population: All cars produced in a day.	3 Marks

4. 100,000 randomly selected adults were asked whether they drink at least 48 oz of water each day and only 45% said yes. Identify the sample and population — 1 Marks

Sample: 100,000 adults surveyed.
Population: All adults in the general population.

5. Match the sampling method to the correct answer: — 4 Marks



a) Simple Random Sampling	A researcher selects every 10th person from a list.
b) Systematic Sampling	A teacher selects students by drawing names from a hat.
c) Stratified Sampling	A survey selects 5 males and 5 females from a class of 50.
d) Cluster Sampling	A country is divided into 5 regions and 2 regions are randomly selected for a study.

6. Explain the difference between population distribution and sampling distribution in table. — 4 Marks

Population Distribution	Sampling Distribution
The distribution of values for all individuals in the population.	The distribution of a statistic (e.g., sample mean) obtained from multiple samples drawn from the population.

7. State three sampling errors and three non-sampling errors. — 6 Marks

Sampling Errors:
- Selection Bias
(Occurs when the sample is not representative of the population due to an improper selection process. For example, surveying only urban residents about national opinions may exclude rural perspectives)
- Sample Size Error
(If the sample size is too small, it may not accurately reflect the characteristics of the entire population, leading to misleading conclusions.)
- Non-Response Bias
(Happens when a significant number of selected participants do not respond, which can skew results if non-respondents have different characteristics from respondents.)

Non-Sampling Errors:
- Measurement Error
(Arises due to faulty data collection methods, such as poorly designed survey questions, inaccurate measuring instruments, or human mistakes in recording data.)
- Processing Error
(Occurs due to errors in data handling, such as incorrect data entry, coding mistakes, or miscalculations during analysis.)
- Response Bias
(Happens when participants provide false or misleading answers due to misunderstanding, social desirability, or reluctance to disclose truthful information.)

8. What is the difference between point estimation and interval estimation? — 3 Marks

Point Estimation	Interval Estimation
Provides a single value as an estimate of the population parameter.	Provides a range of values (confidence interval) that is likely to contain the population parameter.

9. Calculate the point estimation for the population mean and standard deviation from the sample data: 12, 18, 25, 30, 35 — 6 Marks

Mean: $\frac{12+18+25+30+35}{5} = 24$

Standard Deviation:

$$s = \sqrt{\frac{(12-24)^2+(18-24)^2+(25-24)^2+(30-24)^2+(35-24)^2}{5-1}}$$

$$s \approx 9.16$$

10. Define a confidence interval and explain why it is important. — 3 Marks

A confidence interval is a range of values within which the true population parameter is expected to lie with a certain probability (e.g., 95%). It is important because it accounts for sampling variability and provides a measure of precision.

11. Calculate the confidence interval for the population mean if the sample mean is 50, the standard deviation is 10, and the sample size is 36 (Assume 95% confidence level). — 3 Marks

$$CI = \bar{x} \pm Z \times \frac{\sigma}{\sqrt{n}}$$

$$CI = 50 \pm 1.96 \times \frac{10}{\sqrt{36}}$$

$$CI = 50 \pm 3.27$$

$$CI = (46.73, 53.27)$$

12. A sample of size 64 is taken from a normally distributed population with a mean of 200 and a standard deviation of 16. Calculate the probability that the sample mean is between 195 and 205. — 4 Marks

$$Z = \frac{X - \mu}{\sigma/\sqrt{n}}$$

$$Z = \frac{195 - 200}{16/\sqrt{64}} = \frac{-5}{2} = -2.5$$

$$Z = \frac{205 - 200}{16/\sqrt{64}} = \frac{5}{2} = 2.5$$

- Using the standard normal table:
  $P(-2.5 < Z < 2.5) = 0.9876 - 0.0062 = 0.9814$
- **Probability = 98.14%**

13. A researcher collects a sample of size 25 from a population with an unknown standard deviation. The sample mean is 70 and the sample standard deviation is 15. Calculate the 95% confidence interval for the population mean using the t-distribution. $$CI = \bar{x} \pm t \times \frac{s}{\sqrt{n}}$$ • From the t-table, $t_{0.025,24} \approx 2.064$. $$CI = 70 \pm 2.064 \times \frac{15}{\sqrt{25}}$$ $$CI = 70 \pm 6.19$$ $$CI = (63.81, 76.19)$$	4 Marks
14. Calculate the confidence interval for a small sample (n = 9, mean = 25, standard deviation = 5, 95% confidence level) using the t-distribution. • $t_{0.025,8} \approx 2.306$ $$CI = 25 \pm 2.306 \times \frac{5}{\sqrt{9}}$$ $$CI = 25 \pm 3.84$$ $$CI = (21.16, 28.84)$$	7 Marks