Basics of Sampling

Answer: -1

Systematic Sampling for Studying Customer Satisfaction in a Mall:

Systematic sampling is a method where you select every n^{th} individual from a population list after choosing a random starting point. Here's how you could apply it in a mall setting

Step-by-Step Explanation:

1: - Define the population:

- The population is all customers visiting the mall on a given day.

2: -Decide the sample size:

- Suppose you want to survey **100 customers** out of an estimated **1,000 visitors** that day

3: - Calculate the sampling interval (n):

$$n=\frac{\text{Total population}}{\text{Sample size}}=\frac{1000}{100}=10$$

So, you will select every **10th customer**

4: - Choose a random starting point:

- Randomly pick a number between **1 and 10** (say, **4**).
- That means the **4th customer** will be your first participant

5: -Select the sample systematically:

- Then, choose every **10th** customer after that:
- 4th, 14th, 24th, 34th, 44th … and so on until you have 100 participants.

6: - Collect data:

- Approach these selected customers and ask them to fill out a short satisfaction survey.

Advantages:

- Easy to implement and manage in a busy mall.
- Ensures even coverage across the population.
- Reduces bias compared to convenience sampling.

Answer: -2

To estimate the **average exam score** of all students in a school with **5 grades of different sizes**, **stratified sampling** is a smart and statistically sound approach.

Step-by-Step Stratified Sampling Process

1: -Identify the Strata

Each **grade** (e.g., Grade 1 to Grade 5) is treated as a **stratum** because they differ in size and potentially in performance.

2: -Determine the Population Size of Each Stratum

- Count the number of students in each grade. For example:
- Grade 1: 100 students ○ Grade 2: 80 students ○ Grade 3: 120 students ○ Grade 4: 150 students ○ Grade 5: 50 students

3: -Decide on the Total Sample Size

Suppose you want a sample of **100 students** in total.

4: -Allocate Samples Proportionally

- Use **proportional allocation** to ensure each grade is fairly represented based on its size:
- Grade 1: (100/500) × 100 = 20 students ○ Grade 2: (80/500) × 100 = 16 students ○ Grade 3: (120/500) × 100 = 24 students ○ Grade 4: (150/500) × 100 = 30 students ○ Grade 5: (50/500) × 100 = 10 students

5: -Randomly Select Students Within Each Grade

- Use **simple random sampling** or **systematic sampling** within each grade to select the required number of students.

6: - Collect Exam Scores

Administer or collect exam scores from the selected students.

Answer: -3

Step-by-Step Calculation for Proportional Allocation

To ensure proportional representation in stratified sampling, allocate the sample size (100 students) to each group based on its proportion of the total population (800 students). This is done by calculating the proportion of each group and multiplying by the sample size.

1: -Calculate Proportions:

- Junior grades: 400 / 800 = 0.5 (50%)
- Senior grades: 250 / 800 = 0.3125 (31.25%)
- Advanced grades: 150 / 800 = 0.1875 (18.75%)

2: -Multiply by Sample Size:

- Junior: 0.5 × 100 = 50 students
- Senior: 0.3125 × 100 = 31.25 students (round to 31)
- Advanced: 0.1875 × 100 = 18.75 students (round to 19)

Rounding is necessary since sample sizes must be whole numbers. The totals sum to 100 (50 + 31 + 19 = 100).

Recommended Sample Sizes

- **Junior grades**: 50 students
- **Senior grades**: 31 students
- **Advanced grades**: 19 students

This allocation ensures each group is represented in proportion to its size, providing a balanced and unbiased sample for your survey. If variances differ significantly between groups, consider optimal allocation instead, but proportional is standard for equal representation.

Answer: -4

Explanation of Systematic Sampling

In systematic sampling, you select every 20th transaction from the list of 1,000, starting at transaction #7. The sequence is: start at 7, then add 20 each time (7 + 20 = 27, 27 + 20 = 47, etc.).

First 10 Transactions in the Sample

1st: Transaction #7

2nd: Transaction #27

3rd: Transaction #47

4th: Transaction #67

5th: Transaction #87

6th: Transaction #107

7th: Transaction #127

8th: Transaction #147

9th: Transaction #167

10th: Transaction #187

Continue adding 20 for more (e.g., #207 next). This ensures even coverage; if the list ends before 1,000, stop at the last valid transaction.

Answer: -5

Key Benefit:

- **Reduced travel and administrative costs**: By surveying **all households within just 10 randomly selected neighbourhoods**, you avoid the time and expense of reaching scattered households across all 200 neighbourhoods.

Why It Works Well Here:

- Households within a neighbourhood are **geographically close**, making data collection easier.
- It still provides a **representative estimate** if the selected clusters are randomly chosen.

<u>Standard Error Calculation</u>

<u>Answer: -6</u>

To calculate the **standard error of the mean,** use the formula:

<u>SE = Sigma/Sqrt(n)</u>

Sigma(std) = 3 hour

N (sample size) = 36

= sqrt (6)

3/6 = 0.5

Standard error of the mean = 0.5 hours per week.

<u>Answer: -7</u>

To calculate the **standard error of the mean,** use the formula:

<u>SE = Sigma/Sqrt(n)</u>

Sigma(std) = 200 hour

N (sample size) = 50(sample size) sqrt (7.07)

200/7.07 = 28.28 hour

Standard error = 28.28 hours.

<u>Effect of Sample Size on Variability</u>

<u>Answer: -8</u>

To calculate the **standard error of the mean,** use the formula:

SE = Sigma/Sqrt(n)

Sigma = 10

N = 16 sqrt (4),
64 sqrt (8),
256 sqrt (16).

Sigma(std)	Sample size (n)	SE value
10	4	**2.5**
10	8	**1.25**
10	16	**0.625**

<u>Impact of Increasing Sample Size</u>

- As **sample size increases**, the **standard error decreases**.
- This means your sample mean becomes a **more precise estimate** of the population mean.
- **Lower variability** in sample means leads to **more reliable conclusions**.

<u>Conclusion:</u> Larger samples reduce uncertainty and improve accuracy in estimating population parameters.

Real-World Applications

Answer: -9

Sampling Method Used:

This is an example of **voluntary response sampling (also known as self-selection sampling)** — customers choose whether or not to respond to the survey.

Potential Biases:

1. **Non-response Bias** of the 850 customers who didn't respond may have different opinions, leading to skewed results.
2. **Self-selection Bias** ○ Those who responded might feel strongly (positive or negative), while neutral customers may ignore the survey.
3. **Lack of Representativeness**

The sample may not reflect the diversity of the full customer base in terms of demographics, behaviour, or satisfaction.

Answer: -10

Your 10sample means:
75, 78, 74, 76, 77, 75, 79, 76, 74, 77

Sampling distribution of the mean

(Just the list above — those 10 values form the sampling distribution.)

Mean of the sampling distribution

Sum = 761 →

Mean = 761/10 = 76.1

Standard deviation of the sampling distribution

Compute squared deviations from 76.1, sum = **24.90**.

- **Population (true) standard deviation** of these 10 values (divide by 10):
  Variance = 24.90 / 10 = 2.49 → SD = $\sqrt{2.49} \approx$ **1.58**
- **Sample (unbiased) standard deviation** (divide by n−1 = 9):
  Variance = 24.90 / 9 ≈ 2.7667 → SD ≈ **1.66**

Final answers

- Sampling distribution (list): **75, 78, 74, 76, 77, 75, 79, 76, 74, 77**
- Mean = **76.1**
- Standard deviation ≈ **1.58** (or **1.66** if you prefer the unbiased/sample SD)

Interpretation: the standard deviation (~1.6) tells you how much the sample means varies around 76.1 — i.e., an empirical estimate of the sampling variability.

Probability Using CLT

Answer: -11

To find the **probability that the sample mean is greater than 210g**, we use the **standard normal distribution (Z-score)**.

Given:

- Population mean (mu) = 200g
- Standard deviation (sigma) = 50g
- Sample size (n) = 36
- Sample mean threshold = 210g

Step 1: Calculate Standard Error (SE)

SE = Sigma/Sqrt(n)

Sigma = 50

N (sample size) = 36 sqrt (6)

50/6 = ***8.3***

Step 2: Calculate Z-score

XBAR- MU/SE

XBAR (sample mean) = 210g MU

Population mean = 200

SE = STANDARD ERROR = 8.3

(210 – 200 / 8.3) = 1.2

Step 3: Find the probability

From the **standard normal table**:

$$P(Z < 1.20) = 0.8849$$

So,

$$P(Z > 1.20) = 1 - 0.8849 = 0.1151$$

There is about an **11.5% chance** that the sample mean weight of 36 apples will be greater than 210 g.