

# DPP – Sampling & Sampling Distributions

## Sarthak Sahu

**Question 1:** Suppose you want to study customer satisfaction in a mall. Explain how you would use systematic sampling to select participants.

**Answer:** Systematic sampling selects units at regular intervals from an ordered list. Let N be the total number of customers and n be the desired sample size. The sampling interval k is calculated as  $k = N / n$ . A random starting point between 1 and k is chosen, and then every k-th customer is surveyed.

**Question 2:** A school has 5 grades, each with different numbers of students. Describe how you would use stratified sampling to estimate the average exam score of all students.

**Answer:** The student population is divided into strata based on grades. Random samples are selected from each grade proportionally. The overall average exam score is calculated using the formula: Overall Mean =  $\Sigma$  (Stratum Mean  $\times$  Stratum Proportion). This ensures fair representation from all grades.

**Question 3:** A school has 800 students: 400 in junior grades, 250 in senior grades, and 150 in advanced grades. You want to survey 100 students ensuring proportional representation.

**Answer:** Junior:  $(400/800) \times 100 = 50$  students. Senior:  $(250/800) \times 100 \approx 31$  students. Advanced:  $(150/800) \times 100 \approx 19$  students.

**Question 4:** A store wants to survey its customers. It has a list of 1,000 customer transactions and wants to sample every 20th transaction. If the starting point is transaction #7, list the first 10 transactions.

**Answer:** The selected transactions are: 7, 27, 47, 67, 87, 107, 127, 147, 167, and 187.

**Question 5:** A city has 200 neighborhoods. You select 10 neighborhoods at random and survey all households within them. What is the main advantage of cluster sampling?

**Answer:** The main advantage of cluster sampling is cost and time efficiency. Surveying entire neighborhoods reduces travel and administrative effort compared to sampling individual households across the city.

**Question 6:** In a population of students, the average study time is 10 hours per week with a standard deviation of 3 hours. A sample of size 36 is taken.

**Calculate the standard error.**

**Answer:** The standard error is calculated using  $SE = \sigma / \sqrt{n}$ .  $SE = 3 / \sqrt{36} = 0.5$  hours.

**Question 7: If factory light bulb lifespans have a mean of 1,200 hours and a standard deviation of 200 hours, calculate the standard error for samples of size 50.**

**Answer:**  $SE = 200 / \sqrt{50} \approx 28.28$  hours.

**Question 8: A population has a mean of 50 and a standard deviation of 10. Compare the standard error for sample sizes of 16, 64, and 256.**

**Answer:** For  $n = 16$ ,  $SE = 2.5$ . For  $n = 64$ ,  $SE = 1.25$ . For  $n = 256$ ,  $SE = 0.625$ . As sample size increases, variability decreases.

**Question 9: A company sends out a survey to 1,000 customers and receives responses from 150. Identify the sampling method and discuss potential bias.**

**Answer:** The sampling method used is voluntary response sampling. The main bias is non-response bias, as respondents may not represent the entire population.

**Question 10: A researcher collects 10 random samples of size 30. The sample means are 75, 78, 74, 76, 77, 75, 79, 76, 74, and 77. Calculate the mean and standard deviation of the sampling distribution.**

**Answer:** The mean of the sampling distribution is 76.1. The standard deviation of the sampling distribution is approximately 1.7.

**Question 11: The average weight of apples is 200 g with a standard deviation of 50 g. A random sample of 36 apples is taken. What is the probability that the sample mean is greater than 210 g?**

**Answer:** The standard error is  $SE = 50 / \sqrt{36} = 8.33$ . The Z-score is  $(210 - 200) / 8.33 \approx 1.20$ . The probability that the sample mean is greater than 210 g is approximately 0.115 or 11.5%.