CHAPTER 7: SAMPLING DISTRIBUTIONS

The Foundation of Statistical Inference

I. Sampling and Point Estimation

A **Point Estimator** is a sample statistic used to estimate a population parameter.

- Sample mean ($\bar{x}$) estimates population mean (μ).
- Sample standard deviation (s) estimates population std. dev (σ).
- Sample proportion ($\bar{p}$) estimates population proportion (p).

Because different samples yield different results, these estimators are themselves **random variables**.

II. Selecting a Sample

Finite Population: Use a **Simple Random Sample** where every possible sample of size n has the same probability of being selected. **Infinite Population:** A sample is random if each element comes from the same population and is selected independently.

III. Sampling Distribution of $\bar{x}$

The sampling distribution is the probability distribution of all possible values of the sample mean $\bar{x}$.

1. Expected Value

$E(\bar{x}) = \mu$. The sample mean is an **unbiased estimator** of the population mean.

2. Standard Deviation (Standard Error)

Standard deviation of $\bar{x}$ is called the **Standard Error (SE)**.

- **Infinite Population:** $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
- **Finite Population:** $\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}}\left(\frac{\sigma}{\sqrt{n}}\right)$

Teacher's Note: Use the Finite Population Correction (FPC) factor if $n/N > 0.05$.

3. Central Limit Theorem (CLT)

The CLT is the "magic" of statistics. It states that regardless of the shape of the population, the sampling distribution of $\bar{x}$ becomes **Normal** as n increases ($n \geq 30$).

IV. Sampling Distribution of $\bar{p}$

Used for categorical data (e.g.,

1. Expected Value

$E(\bar{p}) = p$. The sample proportion is an unbiased estimator.

2. Standard Error of the Proportion

- **Infinite Population:** $\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$

The distribution of $\bar{p}$ is approximately **Normal** if $np \geq 5$ and $n(1-p) \geq 5$.

V. Practical Implications

As the sample size (n) increases, the Standard Error decreases. This means a larger sample provides a much more **precise** estimate of the population. This is why major political polls use samples of 1,000+ people—to shrink the error.

VI. Step-by-Step Example

Problem: A population has $\mu = 200$ and $\sigma = 50$. If you take a sample of $n = 100$, what is the probability that your sample mean $\bar{x}$ will be within ± 5 of the true population mean?

Logic: 1. **Identify Parameters:** $\mu = 200, \sigma = 50, n = 100$. 2. **Calculate Standard Error:** $\sigma_{\bar{x}} = 50/\sqrt{100} = 50/10 = 5$. 3. **Target Range:** We want $P(195 \leq \bar{x} \leq 205)$. 4. **Calculate Z-scores:**

$$z = \frac{205 - 200}{5} = 1.00 \quad \text{and} \quad z = \frac{195 - 200}{5} = -1.00$$

5. **Find Probability:** Area between $z = -1$ and $z = 1$ is roughly **0.6826** (using Z-table or Empirical Rule).

VII. Practice Set

1. A company has 10,000 employees ($N = 10,000$). You sample 1,000 ($n = 1,000$). Should you use the Finite Population Correction factor?

2. If the population proportion $p = 0.40$ and $n = 100$, calculate the Standard Error of the proportion ($\sigma_{\bar{p}}$).

3. Why does the standard error of the mean decrease as n increases?

4. True/False: The expected value of the sample mean $E(\bar{x})$ is always equal to the population mean μ, regardless of sample size.

VIII. Answer Key

1. **Yes.** $n/N = 1,000/10,000 = 0.10$. Since $0.10 > 0.05$, the FPC is required. 2. $\sigma_{\bar{p}} = \sqrt{(0.4 \times 0.6)/100} = \sqrt{0.0024} = \mathbf{0.049}$. 3. Because we divide σ by $\sqrt{n}$. As the denominator gets larger, the quotient gets smaller. 4. **True.** The sample mean is an unbiased estimator.