6.4 Determine Sample Size for Surveys

Choosing the right sample size is essential for designing a survey. It directly affects the accuracy and reliability of results. A sample that is too small risks failing to capture sufficient diversity and can lead to misleading conclusions, while an overly large sample can waste resources and be inefficient.

Key Factors in Sample Size Calculation

1. **Population Size (N)**
   - Represents the total number of individuals in the population under study. For large populations (e.g., millions), this factor has minimal impact. For smaller populations, it becomes more critical.
2. **Margin of Error (E)**
   - Defines the range within which the true population parameter is likely to fall, expressed as a percentage (e.g., ±5%). Smaller margins of error require larger sample sizes.
   - **Example**: A margin of error of ±5% indicates that the actual value in the population is likely within 5% of the sample estimate.
3. **Confidence Level (Z)**
   - Indicates the degree of certainty desired that the sample represents the population. Common levels are 90%, 95%, and 99%. Higher confidence levels require larger sample sizes.
   - **Z-values**:
     - 90% = 1.645
     - 95% = 1.96
     - 99% = 2.576
4. **Population Proportion (P)**
   - Refers to the expected proportion of the population with a specific characteristic. If unknown, 50% (0.5) is typically used to maximize variability and ensure the largest sample size.

Sample Size Formula for Proportions

For simple random sampling, the required sample size (n) is calculated using:

$$n = \frac{Z^2 \cdot P \cdot (1 - P)}{E^2}$$

Where:
- (Z): Z-value for the desired confidence level.
- (P): Population proportion (0.5 if unknown).
- (E): Margin of error.

Example: Sample Size for Proportions

Parameters:

- Population Size ((N)) = 10,000
- Confidence Level = 95% ((Z = 1.96))
- Estimated Proportion ((P)) = 50% (most conservative estimate)
- Margin of Error ((E)) = ±5% (0.05)

Calculation:

$$n = \frac{(1.96)^2 \cdot 0.5 \cdot (1 - 0.5)}{(0.05)^2}$$

1. Calculate the numerator:

$$1.96^2 \cdot 0.5 \cdot 0.5 = 3.8416 \cdot 0.25 = 0.9604$$

2. Calculate the denominator:

$$(0.05)^2 = 0.0025$$

3. Compute (n):

$$n = \frac{0.9604}{0.0025} = 384.16$$

Thus, the required sample size is approximately **384 participants**.

Sample Size Formula for a Mean

When estimating a population mean, the sample size can be calculated using:

$$n = \frac{Z^2 \cdot \sigma^2}{E^2}$$

Where:
- ( ): Population standard deviation (estimated from prior studies or pilot tests).
- (E): Margin of error.

Example: Sample Size for a Mean

Parameters:

- Confidence Level = 95% ((Z = 1.96))
- Standard Deviation (( )) = 5,000 JDs (estimated)
- Margin of Error ((E)) = ±500 JDs

Calculation:

1. Calculate the numerator:

$$Z^2 \cdot \sigma^2 = 1.96^2 \cdot 5000^2 = 3.8416 \cdot 25,000,000 = 96,040,000$$

2. Calculate the denominator:

$$E^2 = 500^2 = 250,000$$

3. Compute (n):

$$n = \frac{96,040,000}{250,000} = 384.16$$

The required sample size is approximately **384 participants**.

Finite Population Correction (FPC)

For small populations (less than 20,000), adjust the sample size using the FPC formula:

$$n_{\text{adj}} = \frac{n \cdot (N - n)}{N - 1}$$

Where:
- (n): Initial sample size.
- (N): Population size.

Practical Considerations

1. **Budget and Resources**: Larger samples improve precision but increase costs and time. Balance statistical precision with practical constraints.
2. **Sampling Method**: Adjust sample size if using stratified, cluster, or systematic sampling, as these methods can impact efficiency and precision.
3. **Pilot Studies**: Conduct small pilot surveys to refine estimates for parameters like ( ) or (P).

Sample Size for Different Confidence Levels and Margins of Error

Here's a quick reference for sample sizes with a margin of error of ±5%:

Confidence Level	Z-value	Required Sample Size
90%	1.645	271
95%	1.96	384
99%	2.576	664

Conclusion

Determining the appropriate sample size ensures reliable and accurate survey results while optimizing time and resources. Proper calculation considers the desired confidence level, margin of error, and population parameters, ensuring a balance between statistical validity and practicality.