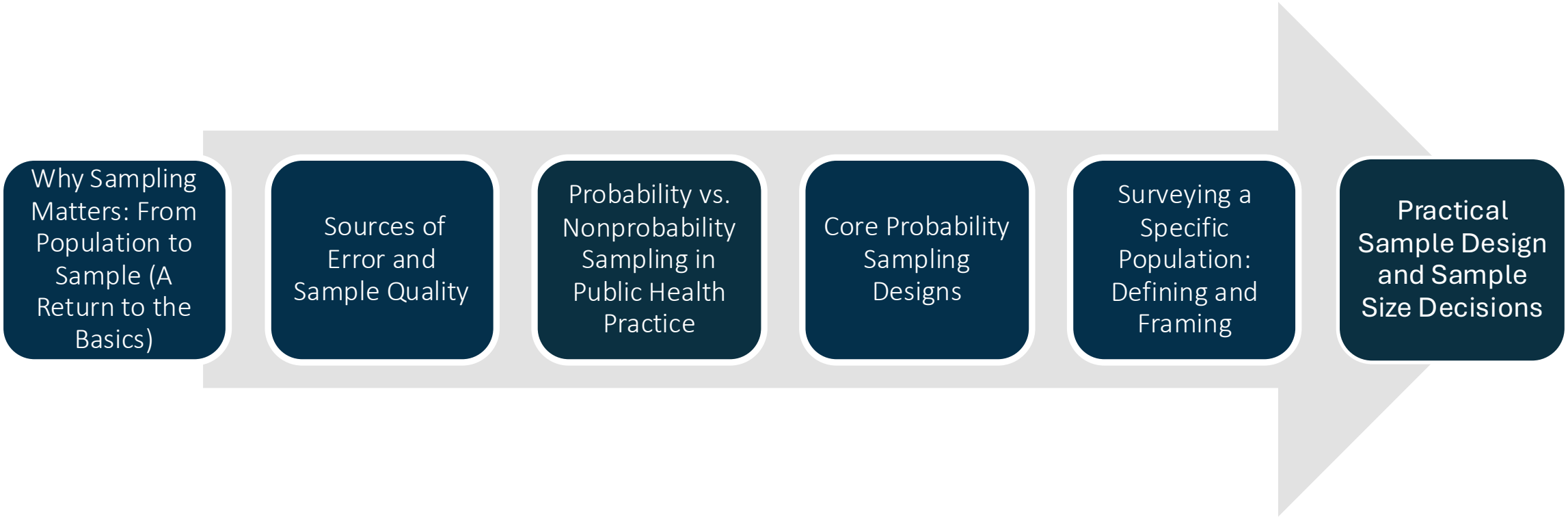
Local Health Department Academy of Science

Survey School

Course Overview



Lesson 1:
Why Sampling Matters:
From Population to Sample

A Return to the Basics

Kevin Brooks, MSc, PhD

Objectives

- Define population, target population, study population, and sample.
- Explain why sampling is used instead of a census in most research settings.
- Describe how sample results are generalized to larger populations.

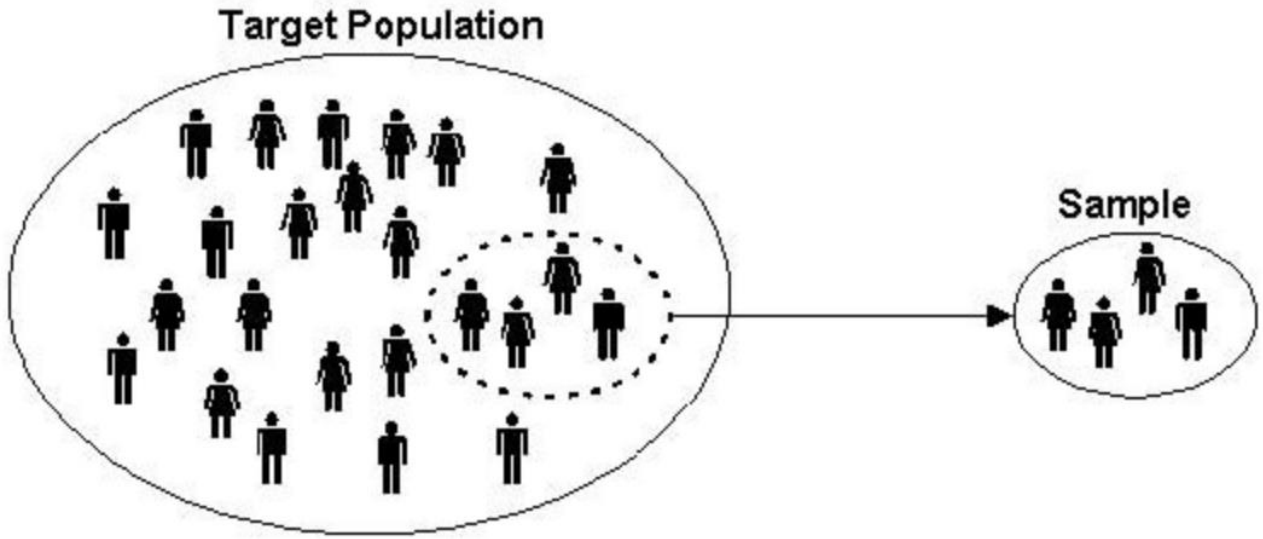
Why Do We Need Sampling?

- Populations are often very large
- Census studies are expensive
- Collecting data from everyone takes time
- Many populations are difficult to reach completely
- Sampling provides a practical alternative

Sampling is one of the most important foundations of survey research

Population, Sample, and Census

Term	Definition	Example
Population	Entire group of interest	All Michigan adults
Sample	Subset selected from the population	2,500 Michigan adults selected for a survey
Census	Information collected from everyone	Every Michigan adult surveyed



Why Most Research Uses Samples

- Lower cost
- Faster results
- Greater feasibility
- Easier data management
- Allows study of large populations

Trade-off

We give up complete coverage in exchange for practical and timely information

What Makes a Good Sample?

- **Nonsampling Error**
  - ✓ Interviewer, response, or coding errors
- **Sampling Error**
  - ✓ Random variation due to studying a sample
- **Sample Bias**
  - ✓ Systematic differences between the sample and population

Important: Bias is usually a larger threat than sample size.

What Makes a Good Sample?

Source	Caused by	Mitigated by
Sampling error (sampling variance)	Random variation in sample composition	Larger sample size and/or efficient sample design
Sample bias		
Coverage bias	Some members of the population are excluded from consideration in the sample	Defining the population in advance and using materials and procedures that maximize coverage
Selection bias	Some members of the population are given disproportionately high or low chances of selection	Using probability sampling and weighting the results for any differences in probabilities of selection
Nonresponse bias	Some members of the sample fail to respond	Using procedures that minimize nonresponse and possibly weighting results to adjust for differences in response rates across groups

The Most Important Lesson in Sampling

- **1936 Literary Digest:**
  - ✓ large sample,
  - ✓ wrong prediction
- **Gallup:**
  - ✓ smaller sample,
  - ✓ correct prediction

Who matters more than how many

Smaller, representative sample usually better than larger, biased sample.

The Most Important Lesson in Sampling

Major shift to area probability after the 1948 presidential elections

All major polls incorrectly predicted victory by Republican challenger Thomas Dewey over Democrat Harry Truman.



Quota Sampling vs. Probability Sampling

Characteristic	Quota Sampling	Probability Sampling
Type of Sampling	Nonprobability	Probability
Selection Method	Respondents selected until quotas are filled; selection within quotas is nonrandom	Respondents selected using random methods
Representativeness	Can resemble population on selected traits only	Designed to represent the entire population
Risk of Selection Bias	Higher	Lower
Cost	Lower	Higher
Common Public Health Uses	Exploratory studies, rapid assessments, pilot surveys	Population health surveys, surveillance systems, official estimates
Example	Recruit adults at health fairs until age and sex quotas are met	Randomly select households or individuals from a sampling frame

Thank you

~ Why Sampling Matters: From Population to Sample

A Return to the Basics ~

Next →Lesson 2. Sources of Error and Sample Quality

Kevin Brooks, MSc, PhD