Marketing Research Test 3 – Chapters 9, 10, 11, 12, 14

Chapter 9: Developing the Sampling Plan

A. Introduction

1. If you collect information from or about each member of the relevant population, you are conducting a **census.** When the overall population is limited in size, it is usually smart to attempt a census, even if you can't get information from everyone.
2. Sampling plan is the process of selecting the people of objects (i.e., companies, products, etc.) to be surveyed, interviewed, or observed.
3. With a sampling plan, we collect information from a **sample,** or subset, of elements from the larger group of objects, with the goal of making projections about what would be true for the population.

B. Defining the Target Population

1. **Population –** all the individuals or objects that meet certain requirements for membership in the overall group
2. We typically work with a sample, rather than a census, because a sample is often easier and less costly to obtain than is a census
3. Any population has certain characteristics; these characteristics are called **parameters**, and we assume that if we could take measurements of these characteristics from all population elements without any error we would know what is true about the population on these parameters.
4. **Statistics –** characteristics or measures of a sample, to draw inferences about the larger population's parameters.
5. **Sampling error –** the difference between results obtained from a sample and results that would have been obtained had information been gathered from or about every member of the population.

C. Identifying the Sampling Frame

1. **Sampling frame –** the list of population elements from which a sample will be drawn; the list could consist of geographic areas, institutions, individuals, or other units
2. Unfortunately, perfect sampling frames usually don't exist except in unusual circumstances. That makes developing an acceptable sampling frame one of your most important and creative tasks.
3. Sometimes you'll work with sampling frames that have been developed by companies that specialize in compiling databases and then selling the names, addresses, phone numbers, and email addresses.
4. Commonly used sampling frames include:
   - Customer databases
   - Telephone directories
   - Lists developed by data compilers

D. Selecting a Sampling Procedure

1. **Nonprobability Samples –** involve personal judgment somewhere in the selection process. Not all elements have an opportunity to be included so we can't estimate the probability that any particular element will be included in the sample. As a result, it's impossible to assess the degree of sampling error. And that means that we can't say anything at all about what would have been true for the overall population – we're stuck with sample statistics and don't know whether they apply to the population as a whole.

2. **Convenience samples –** people or objects are selected for the sample because they happen to be in the right place at the right time to be included. Just go out and find a location where lost of people who are likely to be members of the population are located and do interviews or pass out surveys. Commonly used for with exploratory research, where the goal is to generate insights or to develop hypotheses
   - Also called "accidental" sample – EX: television news "question of the day" polls
3. **Judgment samples –** the sample elements are handpicked by the researcher because she believes that they can serve the research process
   - EX: hire panelists who are knowledgeable about the issue being researched rather than selecting them at random
   - **Snowball sample –** a judgment sample that is used to sample special populations, in particular populations that are difficult to find and identify. This type of judgment sample relies on the researcher's ability to locate an initial set of respondents with the desired characteristics. These individuals are then asked to help identify others with the desired characteristics.
   - **EX**: A demand study for a new product where initial respondents know people with a high interest level within the product category.
4. **Quota samples –** A nonprobability sample chosen so that the proportion of sample elements with certain characteristics is about the same as the proportion of the elements with the characteristics in the target population
   - Stated more simply, certain important characteristics of the population are represented proportionately in the sample
   - EX: Research problem: investigate 100 undergraduate student attitudes toward a controversial new technology fee
   - Known population parameters: class (30% freshman, 20% sophomores, 30% Juniors, 20% seniors) and gender (50% female, 50% male)
   - Approach: 10 students will interview 10 friends each for a total of 100 responses
5. **Probability Samples –** each member of the target population has a *known nonzero* chance of being included in the sample. The chances of each member of the target population being included in the sample may not be equal, but everyone has some chance of being included. Because of this objectivity, we can make inferences to the larger population based on the results from the sample and estimate the likely amount of sampling error
6. **Simple random samples –** each unit included in the sample has a known and equal chance of being selected for study, and every combination of population elements is a sample possibility.
   - Walking down the street and passing out surveys to unknown people "at random" is "random" in the everyday sense, but not random in a scientific sample sense
   - EX: Sample is drawn by a computer or from a physical list using a random number table
7. **Systematic sample –** a probability sampling plan in which every *k*th element in the population is selected for the sample pool after a random start
   - Calculating the **sampling interval** (i.e., *k*, the number of names to count when selecting the same members) – in general, we simply divide the number of population elements in the sampling frame by the number of elements that we need to draw to obtain the sample size we want
   - EX: Example – Research Problem: Investigate 250 undergraduate student attitudes toward controversial new technology fee
   - Known Population: 5000 students published in the campus directory

- Approach: k = 5000/250 = 20 or 1 out of every 20 students on campus will be surveyed. Randomly select the first name then count down 20 names. Select that person to be surveyed and then count down 20 names again. Select that person and so on until you get 250 names.
- **Total Sampling Elements (TSE) –** the number of population elements that must be drawn from the population and included in the initial sample pool in order to end up with the desired sample size
- Calculating the TSE typically requires making predictions about the proportion of sample elements that (1) have incorrect contact information; (2) are ineligible because they don't meet criteria for inclusion in the sample; (3) refuse to participate and (4) cannot be contacted, even after multiple retries

8. **Stratified Samples –** a probability sample in which (1) the population is divided into mutually exclusive and exhaustive subgroups (i.e., each population element fits into one-and only one-subgroup and (2) samples are chosen from each of the subgroups. Most appropriate when subsets (or strata) are homogeneous within but heterogeneous between with respect to key variables
9. Example – Phoenix is one subset, Tucson is a second subset, and all other residents within the state of Arizona constitute a third subset
10. **Cluster samples –** A probability sampling in which (1) the parent population is divided into mutually exclusive and exhaustive subsets, and (2) a random sample of one or more subsets (clusters) is selected – Notice the difference here: with stratified sampling, a sample of elements is selected from each subgroup, but with cluster sampling a sample is selected only from the randomly selected subgroups - Most appropriate when subsets (or strata) are heterogeneous within but homogeneous between with respect to key variables
    - **Area sample –** a form of cluster sampling in which areas (e.g., census tracts, blocks) serve as the primary sampling units. Using maps, the population is divided into mutually exclusive and exhaustive areas, and a random sample of areas is selected.
11. **Nonprobability sample:**
    - Neither sampling error nor the margin of sampling error can be estimated or calculated
    - Inferences cannot be made about the population
    - Inferences are **limited** to the sample
    - Thus, results are not generalizable from the sample to the population
12. **Probability Sample**
    - One can statistically asses level of sampling error
    - Inferences can be made about the population, and not just the sample
    - Inferences are **not limited** to the sample
    - Thus, results are generalizable from the sample to the population

E. Determine the Sample Size

1. Three basic factors affect the size of sample needed when working with a probability sample
    a. Amount of Diversity of Variation of the parameter in question within the population
    - As diversity/variation increases, larger samples are required
    b. Degree of Precision
    - **Precision –** the degree of error in an estimate of a population parameter
    - As need for precision increases, larger samples are required
    c. Degree of confidence
    - **Confidence –** the degree to which one can feel confident than an estimate approximates the true value