CHAPTER 5

Choosing the Type of Probability Sampling

What you will learn in this chapter:

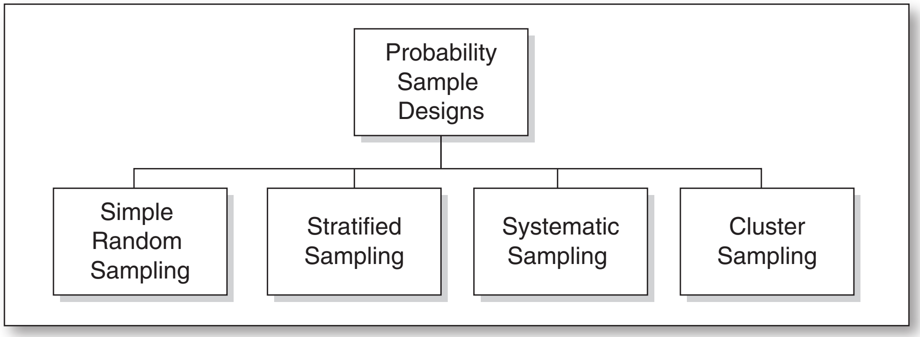
- The types of probability sampling and how they differ from each other
- Steps in carrying out the major probability sample designs
- The strengths and weaknesses of the various types of probability sampling
- Differences between stratified sampling and quota sampling
- Differences between stratified sampling and cluster sampling
- Differences between multistage cluster sampling and multiphase sampling

INTRODUCTION

Once a choice is made to use a probability sample design, one must choose the type of probability sampling to use. This chapter includes descriptions of the major types of probability sampling. It covers steps involved in their administration, their subtypes, their weaknesses and strengths, and guidelines for choosing among them.

There are four major types of probability sample designs: simple random sampling, stratified sampling, systematic sampling, and cluster sampling (see Figure 5.1). Simple random sampling is the most recognized probability sampling procedure. Stratified sampling offers significant improvement to simple random sampling. Systematic sampling is probably the easiest one to use, and cluster sampling is most practical for large national surveys. These sampling procedures are described below.

Figure 5.1 Major Types of Probability Sampling



SIMPLE RANDOM SAMPLING

What Is Simple Random Sampling?

Simple random sampling is a probability sampling procedure that gives every element in the target population, and each possible sample of a given size, an equal chance of being selected. As such, it is an **equal probability selection method** (EPSEM).

What Are the Steps in Selecting a Simple Random Sample?

There are six major steps in selecting a simple random sample:

1. Define the target population.
2. Identify an existing sampling frame of the target population or develop a new one.
3. Evaluate the sampling frame for undercoverage, overcoverage, multiple coverage, and clustering, and make adjustments where necessary.
4. Assign a unique number to each element in the frame.
5. Determine the sample size.
6. Randomly select the targeted number of population elements.

Three techniques are typically used in carrying out Step 6: the lottery method, a table of random numbers, and randomly generated numbers using a computer program (i.e., random number generator). In using the lottery method (also referred to as the "blind draw method" and the "hat model"), the numbers representing each element in the target population are placed on chips (i.e., cards, paper, or some other objects). The chips are then placed in a container and thoroughly mixed. Next, blindly select chips from the container until the desired sample size has been obtained. Disadvantages of this method of selecting the sample are that it is time-consuming, and is limited to small populations.

A table of random numbers may also be used. The numbers in a table of random numbers are not arranged in any particular pattern. They may be read in any manner, i.e., horizontally, vertically, diagonally, forward, or backward. In using a table of random numbers, the researcher should blindly select a starting point and then systematically proceed down (or up) the columns of numbers in the table. The number of digits that are used should correspond to the total size of the target population. Every element whose assigned number matches a number the researcher comes across is selected for the sample. Numbers the researcher comes across that do not match the numbers assigned the elements in the target population are ignored. As in using the lottery method, using a table of random numbers is a tedious, time-consuming process, and is not recommended for large populations. Instead, statistical software should be used for large populations. Most statistical software and spreadsheet software have routines for generating random numbers. Elements of the populations whose assigned numbers match the numbers generated by the software are included in the sample. One may select a number from a table of random numbers for use as the starting number for the process.

What Are the Subtypes of Simple Random Sampling?

There are two types of simple random sampling: sampling with replacement and sampling without replacement. In sampling with replacement, after an element has been selected from the sampling frame, it is returned to the frame and is eligible to be selected again. In sampling without replacement, after an element is selected from the sampling frame, it is removed from the population and is not returned to the sampling frame. Sampling without replacement tends to be more efficient than sampling with replacement in producing representative samples. It does not allow the same population element to enter the sample more than once. Sampling without replacement is more common than sampling with replacement. It is the type that is the subject of this text.

What Are the Strengths and Weaknesses of Simple Random Sampling?

Simple random sampling has the major strengths and weaknesses of probability sampling procedures when compared to nonprobability sampling procedures. Notably, among its strengths, it tends to yield representative samples, and allows the use of inferential statistics in analyzing the data collected. Compared to other probability sampling procedures, simple random sampling has several strengths that should be considered in choosing the type of probability sample design to use (see Table 5.1). Some of these include:

- Advanced auxiliary information on the elements in the population is not required. Such information is required for other probability sampling procedures, such as stratified sampling.
- Each selection is independent of other selections, and every possible combination of sampling units has an equal and independent chance of being selected. In systematic sampling, the chances of being selected are not independent of each other.
- It is generally easier than other probability sampling procedures (such as multistage cluster sampling) to understand and communicate to others.
- Statistical procedures required to analyze data and compute errors are easier than those required of other probability sampling procedures.
- Statistical procedures for computing inferential statistics are incorporated in most statistical software and described in most elementary statistics textbooks.

On the other hand, simple random sampling has important weaknesses. Compared to other probability sampling procedures, simple random samplings have the following weaknesses:

- A sampling frame of elements in the target population is required. An appropriate sampling frame may not exist for the population that is targeted, and it may not be feasible or practical to construct one. Alternative sampling procedures, such as cluster sampling, do not require a sampling frame of the elements of the target population.
- Simple random sampling tends to have larger sampling errors and less precision than stratified samples of the same sample size.

- Respondents may be widely dispersed; hence, data collection costs might be higher than those for other probability sample designs such as cluster sampling.
- Simple random sampling may not yield sufficient numbers of elements in small subgroups. This would not make simple random sampling a good choice for studies requiring comparative analysis of small categories of a population with much larger categories of the population.

Research Note 5.1 below describes simple random sampling procedures used in a study of inmate-on-inmate sexual assaults in California's prisons.

Table 5.1 Strengths and Weaknesses of Simple Random Sampling Compared to Other Probability Sampling Procedures

Strengths	**Weaknesses**
Compared to other probability sampling procedures:	Compared to other probability sampling procedures:
Advanced auxiliary information on the elements in the population is not required.	A sampling frame of elements in the target population is required.
Every possible combination of sampling units has an equal and independent chance of being selected.	Does not take advantage of knowledge of the population that the researcher might have.
Easier to understand and communicate to others.	May have larger sampling errors and less precision, than other probability sampling designs with the same sample size.
Tends to yield representative samples.	If subgroups of the population are of particular interests, they may not be included in sufficient numbers in the sample.
Statistical procedures required to analyze data and compute errors are easier.	If the population is widely dispersed, data collection costs might be higher than those of other probability sample designs.
Statistical procedures for computing inferential are incorporated in most statistical software.	May be very costly, particularly where populations are geographically dispersed and/or individuals may be difficult to locate because of change of last name due to marriage or migration.

RESEARCH NOTE 5.1

Example of Simple Random Sampling: Study of Inmate-on-Inmate Sexual Assaults in California's Prisons

Jenness, Maxson, Sumner, and Matsuda (2010) conducted a survey of adult prisoners in California's prisons in a study of inmate-on-inmate sexual assaults. Six prisons were selected using purposive sampling, and then simple random sampling was used to select inmates from the selected prisons. The authors described their simple random sampling procedures as follows:

> We relied on a similar process in each facility to randomly sample inmates. About a week prior to the first day of data collection at a particular prison, the CDCR [California Department of Corrections and Rehabilitation] Office of Research sent us a facility roster that identified every inmate housed in the prison. The roster indicated the inmate's name, CDC number, custody level, classification score, housing location in the facility, and mental health status. Inmates housed in reception centers were excluded. Once we received the roster, we removed inmates categorized as EOP [Enhanced Outpatient, mental patients indicating the highest level of mental incapacity]. Importantly, inmates with other mental health designations (e.g., Correctional Clinical Case Management System [CCCMS]) and inmates on restricted status (e.g., inmates housed in administrative segregation or security housing units [SHUs]) were retained on the final roster from which we randomly selected study participants.
>
> From the final roster, we used statistical software to randomly select 100 inmates from each prison to be study participants. This approach ensured that CDCR officials could not interfere with the random selection on purpose or inadvertently. We randomly ordered the CDCR numbers of selected study participants to eliminate bias and sent the list of selected inmates to our liaison at the prison, typically the Public Information Officer or another Lieutenant, so that inmates were scheduled and notified by a ducat to meet with an interviewer on the research team. Maintaining consistency in our sampling procedures, including providing detailed written instructions to our liaisons, supported our goal of attaining a representative sample of inmates.

Source: Jenness, Maxson, Sumner, & Matsuda, 2010 pp. 11-12. Reprinted with permission.

STRATIFIED SAMPLING

What Is Stratified Sampling?

Stratified sampling is a probability sampling procedure in which the target population is first separated into mutually exclusive, homogeneous segments (strata), and then a simple random sample is selected from each segment (stratum). The samples selected from the various strata are then combined into a single sample. This sampling procedure is sometimes referred to as "quota random sampling."

What Are the Steps in Selecting a Stratified Sample?

There are eight major steps in selecting a stratified random sample:

1. Define the target population.
2. Identify stratification variable(s) and determine the number of strata to be used. The stratification variables should relate to the purposes of the study. If the purpose of the study is to make subgroup estimates, the stratification variables should be related to those subgroups. The availability of auxiliary information often determines the stratification variables that are used. More than one stratification variable may be used. However, in order to provide expected benefits, they should relate to the variables of interest in the study and be independent of each other. Considering that as the number of stratification variables increases, the likelihood increases that some of the variables will cancel the effects of other variables, not more than four to six stratification variables and not more than six strata for a particular variable should be used.
3. Identify an existing sampling frame or develop a sampling frame that includes information on the stratification variable(s) for each element in the target population. If the sampling frame does not include information on the stratification variables, stratification would not be possible.
4. Evaluate the sampling frame for undercoverage, overcoverage, multiple coverage, and clustering, and make adjustments where necessary.
5. Divide the sampling frame into strata, categories of the stratification variable(s), creating a sampling frame for each stratum. Within-stratum

differences should be minimized, and between-strata differences should be maximized. The strata should not be overlapping, and altogether, should constitute the entire population. The strata should be independent and mutually exclusive subsets of the population. Every element of the population must be in one and only one stratum.

6. Assign a unique number to each element.
7. Determine the sample size for each stratum. The numerical distribution of the sampled elements across the various strata determines the type of stratified sampling that is implemented. It may be a proportionate stratified sampling or one of the various types of disproportionate stratified sampling.
8. Randomly select the targeted number of elements from each stratum. At least one element must be selected from each stratum for representation in the sample; and at least two elements must be chosen from each stratum for the calculation of the margin of error of estimates computed from the data collected.

What Are the Subtypes of Stratified Sampling?

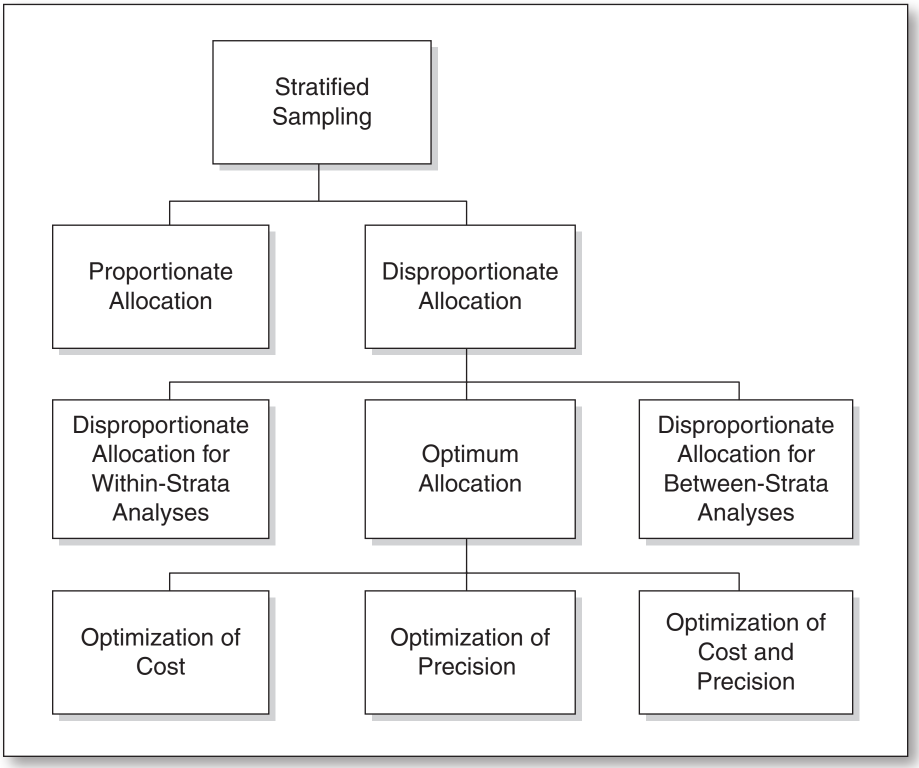
There are two major subtypes of stratified sampling: proportionate stratified sampling and disproportionate stratified sampling (see Figure 5.2). Disproportionate stratified sampling has various subcategories.

Proportionate Stratified Sampling

In proportionate stratified sampling, the number of elements allocated to the various strata is proportional to the representation of the strata in the target population. That is, the size of the sample drawn from each stratum is proportional to the relative size of that stratum in the target population. As such, it is a self-weighting and EPSEM sampling procedure. The same sampling fraction is applied to each stratum, giving every element in the population an equal chance to be selected. The resulting sample is a self-weighting sample. This sampling procedure is used when the purpose of the research is to estimate a population's parameters.

A hypothetical example of **proportionate allocation** is presented in Table 5.2. In this example, the elements sampled were allocated across the four districts of a marketing region so as the proportion of elements sampled for each district is identical to the proportion of elements in each district in the total

Figure 5.2 Subtypes of Stratified Sampling Based on Stratum Allocation



population. The sampling fraction in each district is the same 1 out of 22 elements. Each district is equally represented in the sample.

At times, a researcher may not only desire to estimate population parameters but also to make detailed analyses within a relatively small stratum and/or compare strata to each other. Proportionate stratified sampling may not yield sufficient numbers of cases in some of the strata for such analyses. Taking the example described in Table 5.2 as an example, it would not be possible to conduct a detailed analysis of elements in District 2 because only 12 elements are in the sample. Moreover, comparing District 2 elements to the elements in the other districts would be suspect. Proportionate stratified sampling is a poor sampling choice for carrying out such analyses. Disproportionate stratified sampling may be a better choice.

Table 5.2 Example of Proportionate Stratified Sampling

Marketing Region	Population		Proportionate Stratified Sample	
	Frequency	Percent	Frequency	Percent
District 1	18000	33%	396	33%
District 2	600	1%	12	1%
District 3	12000	22%	264	22%
District 4	24000	44%	528	44%
Total	54600	100%	1200	100%

Disproportionate Stratified Sampling

Disproportionate stratified sampling is a stratified sampling procedure in which the number of elements sampled from each stratum is not proportional to their representation in the total population. Population elements are not given an equal chance to be included in the sample. The same sampling fraction is not applied to each stratum. On the other hand, the strata have different sampling fractions, and as such, this sampling procedure is not an EPSEM sampling procedure. In order to estimate population parameters, the population composition must be used as weights to compensate for the disproportionality in the sample. However, for some research projects, disproportionate stratified sampling may be more appropriate than proportionate stratified sampling.

Disproportionate stratified sampling may be broken into three subtypes based on the purpose of allocation that is implemented. The purpose of the allocation could be to facilitate within-strata analyses, between-strata analyses, or optimum allocation. Optimum allocation may focus on the optimization of costs, the optimization of precision, or the optimization of both precision and costs.

Disproportionate allocation for within strata analyses. The purpose of a study may require a researcher to conduct detailed analyses within the strata of the sample. If using proportionate stratification, the sample size of a stratum is very small; it may be difficult to meet the objectives of the study. Proportionate allocation may not yield a sufficient number of cases for such detailed analyses. One option is to oversample the small or rare strata. Such oversampling would create a disproportional distribution of the strata in the sample when compared to the population. Yet, there may be a sufficient number of cases to carry out the within-strata analyses required by the study's objectives. Examples of research for which such a sample design would be appropriate include a study of Muslims in the military, a study of persons with a rare medical problem, or a study of persons who spent most of their youth in foster care. Using the hypothetical example described in Table 5.2, if it was desired to conduct a detailed analysis of District 2, one might oversample elements from that district; for example, instead of sampling only 12 elements, sample 130 elements. In order to conduct a meaningful, detailed analysis within District 2, the sample size for that district must be larger than 12 elements. The resulting distribution of elements in the sample by district may look like the distribution presented in Table 5.3.

Disproportionate allocation for between-strata analyses. The purpose of a study may require a researcher to compare strata to each other. If this is the case, sufficient numbers of elements must be selected for each category. A researcher may desire to maximize the sample size of each stratum. For such a study, equal allocation (also referred to as "balanced allocation" and "factorial sampling") may be appropriate. A researcher may seek to select an equal number of elements from each stratum.

An example of equal allocation disproportionate allocation is presented in Table 5.4. In this example, the elements sampled were allocated across the four districts of our hypothetical example so that the number of elements sampled for each district is equal. Compared to the proportionate sample distribution in Table 5.2, the sampling allocations presented in Table 5.4 provide a minimum number of elements for each district, making for a more

Table 5.3 Example of Disproportionate Allocation Stratified Sampling

Marketing Region	Population		Disproportionate Stratified Sample	
	Frequency	Percent	Frequency	Percent
District 1	18000	33%	357	30%
District 2	600	1%	130	11%
District 3	12000	22%	238	20%
District 4	24000	44%	475	39%
Total	54600	100%	1200	100%

Table 5.4 Example of Disproportionate Allocation Stratified Sampling to Facilitate Between-Strata Analyses

Marketing Region	Population		Disproportionate Stratified Sample Using Equal Allocation	
	Frequency	Percent	Frequency	Percent
District 1	18000	33%	300	25%
District 2	600	1%	300	25%
District 3	12000	22%	300	25%
District 4	24000	44%	300	25%
Total	54600	100%	1200	100%

balanced comparative analysis across the districts. Moreover, a relatively large number of elements are sampled from District 2, permitting detailed analysis within that stratum.

Optimum allocation. Although proportionate stratified sampling may yield smaller margins of error than simple random sampling in estimating population parameters, it may be possible to do better yet. **Optimum allocation** is designed to achieve even greater overall accuracy than that achieved using proportionate stratified sampling. It sets the sample size of the different strata, taking into account two important aspects of doing research: costs and precision. The sampling fraction varies according to the costs and variability within the various strata. Disproportionate stratified sampling, more specifically, optimum allocation, may be more appropriate for a study than proportionate stratified sampling when the strata differ in terms of data collection costs and the variability of the variables of interest. Optimum allocation may be applied focusing on cost only, precision only, or both cost and precision jointly.

Homogeneous strata with a smaller sample size can have the same level of precision as heterogeneous strata with a larger sample size. Applying this principle, it may be useful to make the number of elements selected from each stratum directly related to the standard deviation of the variable of interest in the stratum. The greater the variability of the variable in a stratum, the higher the sample size of the stratum should be. Moreover, taking into account data collection costs, the higher the data collection costs of a stratum, the lower the targeted sample size. Hypothetical data illustrating optimum allocation are presented in Table 5.5.

The hypothetical data presented in Table 5.5 indicate that the data collection costs within the four districts range from $10 to $39 per unit (see Column 4). Differences in the geographical distribution of the elements in the different strata may account for these differences. The distribution of sample sizes for the various strata in Column 7 takes into account these varying data collection costs.

Hypothetical standard deviations of the variable of interest for the four districts are presented in Column 5 of Table 5.5. The standard deviations range from 4.3 to 9.4. If data collection costs for the various districts are unavailable or essentially the same, one may yet optimize the sample sizes of the various strata by allocating the sample size of each stratum by taking into account the variability of the strata. This type of allocation was first proposed by Jerzy Neyman (1934), and is often referred to as the Neyman allocation. A distribution of the sample size of the different strata in the hypothetical example, taking into account the variability of the strata, is presented in Column 8. The use of this optimization procedure is dependent on data on the variability of the variable of interests for the different strata. Often such data are not available. Moreover, if the study has multiple purposes and more than one variable of interest, their optimization might conflict with each other.

Table 5.5 Examples of Optimum Allocation Disproportionate Stratified Sampling

Marketing Regions (1)	Population Frequency (2)	Population Percent Distribution (3)	Data Collection Cost Per Unit (j) (4)	Variability (s) (5)	$\frac{s}{\sqrt{j}}$ (6)	Sample Size Optimizing Costs (7)	Sample Size Optimizing Variability (8)	Sample Size Optimizing Costs and Variability (9)
District 1	18000	33%	$18	4.3	1.014	300	190	203
District 2	600	1%	$10	6.4	2.024	538	282	405
District 3	12000	22%	$39	9.4	1.505	138	415	302
District 4	24000	44%	$24	7.1	1.449	224	313	290
Total	54600	100%				1200	1200	1200

If data are available for both the data collection costs and the variability of the variable of interest, one may optimize for both costs and precision. A weighting factor taking into account both data collection costs and standard deviation may be computed as $s/\sqrt{j}$, where "s" represents the standard deviation within the stratum, and "j" represents the per-unit data collection costs within the strata. A distribution of this factor is presented in Column 6 of Table 5.5 for our hypothetical example. Taking this factor into account, the sample sizes for the various strata were optimized, taking into account both the data collection costs and the variability with the strata and presented in Column 9.

What Are the Strengths and Weaknesses of Stratified Sampling?

Stratified sampling has many of the strengths and weaknesses associated with most probability sampling procedures when they are compared to non-probability sampling procedures. In determining whether to choose stratified sampling, one may compare its strengths and weaknesses to those of simple random sampling (see Table 5.6). Compared to simple random sampling, the strengths of stratified sampling include:

- Ability to not only estimate population parameters, but also to make within-stratum inferences and comparisons across strata. Sufficient data on subgroups of interest may not be captured in simple random sampling. Stratified samples yield smaller random sampling errors than those obtained with a simple random sample of the same sample size, especially if optimum allocation is used. Stratification makes for a gain in precision, eliminating the variation of the variable that is used for stratifying. The amount of gain in precision is determined by the extent the within-stratum variances of the study variables are minimized and the between-stratum variances of the study variables are maximized. Stratification will yield a sample that is at least as precise as a simple random sample of the same sample size.
- Stratified samples yield smaller random sampling errors than those obtained with a simple random sample of the same sample size, especially if optimum allocation is used. Stratification makes for a gain in precision, eliminating the variation of the variable that is used for stratifying. The amount of gain in precision is determined by the extent the within-stratum variances of the study variables are minimized and the between-stratum variances of the study variables are maximized. Stratification will yield a sample that is at least as precise as a simple random sample of the same

size. If it is ineffective in increasing the level of precision, the results would not be worse than if simple random sampling were used.

- Stratified samples tend to be more representative of a population because they ensure that elements from each stratum in the population are represented in the sample. Sampling may be stratified to ensure that the sample is spread over geographic subareas and population subgroups.
- In using stratified sampling, advantage is taken of knowledge the researcher has about the population.
- If the stratification variable breaks up the population into homogeneous geographical areas, data collection costs may be lower than the data collection costs of sample random sampling.
- Utilizing stratified sampling permits the researcher to use different sampling procedures within the different strata.
- In using stratified sampling, a researcher may be created taking into account administrative convenience in carrying out the study. The researcher may take into account the clustering of the population in metropolitan areas, institutionalized segments of the population, and the distribution of data collection staff.

Compared to simple random sampling, weaknesses of stratified sampling include:

- Stratified sampling has a greater requirement for prior auxiliary information than is the case for simple random sampling. Information on stratification variables is required for each element in the population. Such information includes information on the proportion of the target population that belongs to each stratum; if optimum allocation is used, information on the variability of the variables of interest and information on the data collection costs are necessary for each stratum. Acquiring such information may be time-consuming and costly.
- Selection of stratification variables may be difficult if a study involves a large number of variables. These variables should be correlated with the variables of interests in the study.
- Stratified sampling requires more effort in terms of preparation for sampling, executing the sample design, and analyzing the data collected.
- In order to calculate sampling estimates, at least two elements must be selected from each stratum.
- The analysis of data collected is more complex than the analysis of data collected via simple random sampling.
- Misclassification of elements into strata may increase variability.

- If disproportionate allocation is used, the data collected must be adjusted (weighted) in estimating population parameters. The effect of the weighting is to lower precision of some population estimates.

Table 5.6 Strengths and Weaknesses of Stratified Sampling Compared to Simple Random Sampling

Strengths	Weaknesses
Unlike simple random sampling, stratified sampling:	**Unlike simple random sampling, stratified sampling:**
Has greater ability to make inferences within a stratum and comparisons across strata.	Requires information on the proportion of the total population that belongs to each stratum.
Has slightly smaller random sampling errors for samples of same sample size, thereby requiring smaller sample sizes for the same margin of error.	Information on stratification variables is required for each element in the population. If such information is not readily available, they may be costly to compile.
Obtains a more representative sample because it ensures that elements from each stratum are represented in the sample.	More expensive, time-consuming, and complicated than simple random sampling.
Takes greater advantage of knowledge the researcher has about the population.	Selection of stratification variables may be difficult if a study involves a large number of variables.
Data collection costs may be lower if the stratification variable breaks up the population into homogeneous geographical areas, or so as to facilitate data collection.	In order to calculate sampling estimates, at least two elements must be taken in each stratum.
Permits different research methods and procedures to be used in different strata.	The analysis of the data collected is more complex than the analysis of data collected via simple random sampling.
Permits analyses of within-stratum patterns and separate reporting of the results for each stratum.	If disproportionate allocation is used, weighting is required to make accurate estimates of population parameters.

What Is the Difference Between Stratified Sampling and Quota Sampling?

Stratified sampling and quota sampling are somewhat similar to each other. Both involve dividing the target population into categories and then selecting a certain number of elements from each category (see Table 5.7). Both procedures have as a primary purpose the selection of a representative sample and/or the facilitation of subgroup analyses. However, there are important differences. Stratified sampling utilizes a simple random sampling once the categories are created; quota sampling utilizes availability sampling. A sampling frame is required for stratified sampling, but not for quota sampling. More importantly, stratified sampling is a probability sampling procedure permitting the estimation of sampling error. This is not possible with quota samples.

Listed below are research notes presenting examples of stratified sampling. Research Note 5.2 describes a proportionate allocation stratified sample of students at a Southern university in a study of perception of racism. The next three research notes provide examples of disproportionate allocation stratified samples. Research Note 5.3 describes a disproportionate stratified sample design used in a study of police chiefs. The sample was stratified by size of the city. Large cities and cities with Latino chiefs were oversampled. Research Note 5.4 describes a study of HIV risk behavior among prison inmates. In order to have a sufficient number of women in their study, disproportionate allocation was used oversampling female inmates. The research described in Research Note 5.5 examines differences in the relationship between socioeconomic status and

Table 5.7 Comparison of Stratified Sampling and Quota Sampling

Stratified Sampling	Quota Sampling
Stratified sampling and quota sampling are similar in that:	
Population is divided into categories; elements are then selected from each category.	Population is divided into categories; elements are then selected from each category.
Purpose is to select a representative sample and/or facilitate subgroup analyses.	Purpose is to select a representative sample and/or facilitate subgroup analyses.
Stratified sampling and quota sampling are dissimilar in that:	
Elements within each category are selected using simple random sampling, and as a result:	Elements within each category are selected using availability sampling, and as a result:
A sampling frame is required.	A sampling frame is not required.
Random sampling error can be estimated.	Random sampling error cannot be estimated.
Selection bias is minimized.	Selection bias is not minimized.
Purpose is to reduce sampling error.	

health among African Americans and Whites. In order to have a sufficient number of African Americans, they were oversampling via a disproportionate stratified sampling design.

RESEARCH NOTE 5.2

Example of Proportionate Stratified Sampling: Study of Perception of Racism Among Students at a Southern University

Marcus et al. (2003) utilized proportionate stratified sampling in their study of students' perceptions of racial discrimination in classrooms, on campus, and in contacts with instructors at a Southern university. They described their sampling as follows:

> The data for this study were collected from 398 students who were in 26 randomly selected classes during the spring quarter of 1998. The 26 classes were selected from the entire 555 class sections, excluding laboratory sections and internships, from all of the academic schools using a proportionate stratified sampling approach. Classes in all periods of the day, night, and the weekend were included.
>
> The proportionate sampling plan called for 60% of the sample from day classes, 35% of the sample from evening classes, and 5% of the sample from weekend classes. (This accurately represents the proportion of these classes in these time periods.) This approach resulted in the 26 selected classes; 16 day classes (62%), 9 evening classes (35%) and one weekend class (4%). Of the original 26 classes selected seven instructors (27%) refused permission for their classes to participate. An additional seven classes were selected taking into account the time of day/week of the class. One (14%) of these seven classes was not allowed to participate. It, too, was replaced.
>
> The class rolls of the 26 selected classes indicated a total of 496 enrolled students. The 398 returned instruments are 80% of the enrolled students in the selected classes. If we consider that on the day of administration as many as 10% of the students were either not in attendance or had dropped the course, then the rate of return was over 90%. On the day of administration, 100% of those in attendance responded.

Source: Marcus et al., 2003, p. 614. Reprinted with permission.

RESEARCH NOTE 5.3

Example of Disproportionate Stratified Sampling: Study of Police Chiefs

Hays, Regoli, and Hewitt (2007) used a two-stage stratified sample in their study of police chiefs. They described their sampling procedures as follows:

> Data were derived from a national sample of 1,500 American police chiefs. . . . The sampling frame was constructed from each state's Chiefs of Police Association, which provided separate lists. A two-stage random stratified sampling procedure was used to select participants. First, departments were coded by city size, and eight categories were created to obtain a sample with representatives from all size cities. Second, 200 chiefs were chosen from the first six categories: less than 3,000 in population; 3,000 to 4,999; 5,000 to 9,999; 10,000 to 24,999; 25,000 to 49,999; and 50,000 to 99,999. Because there were only few cities with more than 100,000 populations, all the cases were chosen from the largest two categories, 100,000 to 499,999 and more than 500,000.
>
> In addition to oversampling large cities, Latino chiefs were also oversampled to obtain a sufficient number of Latino chiefs. This was accomplished by carefully perusing the entire sampling frame (more than 10,000 names) and identifying the surnames that appeared to be Latino. Although this methodology is not without its shortcomings, it did result in 77 self-identified Latino chiefs.

Source: Hays, Regoli, & Hewitt, 2007, pp. 8–9. Reprinted with permission.

RESEARCH NOTE 5.4

Example of Disproportionate Stratified Sampling: Study of HIV Risk Behavior Among Prison Inmates

Swartz, Lurigio, and Weiner (2004) used disproportionate sampling in assessing Illinois prison inmates' sexual and drug-use practices, their knowledge about HIV risk-reduction techniques, and their beliefs regarding their own HIV-risk status

(Continued)

(Continued)

and their ability to avoid HIV infection. They described their sampling procedure as follows:

> Research staff recruited participants from the four reception and classification centers (RCCs) that process admittees to the Illinois Department of Corrections (IDOC) prisons in Joliet, Graham, Dwight, and Menard. All IDOC admittees, 18 years or older, were eligible for the study, with the exceptions of federal prisoners, inmates admitted to boot camps, those sentenced to death row, and those sent to RCCs for a transfer to other facilities. To select recruits for the study, we used a probability sampling strategy based on the proportionate number of admissions to each RCC. In addition, because women constituted a small proportion of IDOC admissions (about 7%), we oversampled them, relative to men, at an approximate ratio of 2 to 1. . . . Interviewers sampled participants on site during each day of interviewing. Because of the large variation in the numbers of inmates processed at each RCC, the sampling strategy varied across the four sites. Joliet Prison processed the largest number of admissions. Using a table of random numbers, interviewers selected a sample of inmates to be recruited for the study on that day only. At Graham Prison, interviewers randomly selected every second inmate for study recruitment. At Dwight and Menard Prisons, because of the small number of inmates processed there each week (e.g., on average, Dwight Prison processed approximately 30 to 40 inmates per week), interviewers selected all processed inmates for study recruitment.

Source: Swartz, Lurigio, & Weiner, 2004, pp. 491–492. Reprinted with permission.

RESEARCH NOTE 5.5

Example of Disproportionate Stratified Sampling: Study of the Relationship of Socioeconomic Status and Health Among African Americans and Whites

Ostrove, Feldman, and Adler (1999) analyzed the differences between African Americans and Whites in the relationship between socioeconomic status and health. They used data from two nationally representative surveys of adults in the United States: the Americans' Changing Lives (ACL) survey and the Health and

Retirement Survey (HRS). These samples were appropriate for their study as they were stratified samples with disproportionate allocation due to an oversampling of African Americans. They described the sample designs as follows:

> The ACL survey is a national longitudinal panel survey of African-American and white non-institutionalized adults that was designed to investigate adult activities and social relationships, and adaptation to life events and stress. The first wave of data collection in 1986 used a multistage stratified area probability sampling strategy, with oversampling of African-Americans and those over 60 years of age, and obtained responses from 3617 people. . . . The data were weighted to adjust for variations in probabilities of selection and in response rates, making the data representative of the US population. . . . The HRS is a national panel survey of non-institutionalized adults between the ages of 51 and 61 years (in 1992) and their spouses. The data for the current study are from the original wave of data collection from 1992, in which over 7600 households were sampled, yielding interviews with over 12,600 people. . . . The study used a multistage area probability sample design and oversampled for African-Americans, Latino/as, and residents of Florida. . . . The data were weighted to adjust for unequal selection probabilities, and for geographic and race group differences in response rates, creating a nationally representative sample.

Source: Ostrove, Feldman, & Adler, 1999, p. 454. Reprinted with permission.

SYSTEMATIC SAMPLING

What Is Systematic Sampling?

Systematic sampling (or interval random sampling) is a probability sampling procedure in which a random selection is made of the first element for the sample, and then subsequent elements are selected using a fixed or systematic interval until the desired sample size is reached. The random start distinguishes this sampling procedure from its nonprobability counterpart, nonprobability systematic sampling (discussed above). In some instances, a sampling frame is not used. The target population need not be numbered and a sampling frame compiled if there is physical presentation such as a continuous flow of population elements at specific locations. For example, after a random start, one may systematically select every *i*th patient visiting an emergency room in a hospital, store customers standing in line, or records in file drawers.

What Are the Steps in Selecting a Systematic Sample?

Generally, there are eight major steps in selecting a systematic sample:

1. Define the target population.
2. Determine the desired sample size (n).
3. Identify an existing sampling frame or develop a sampling frame of the target population.
4. Evaluate the sampling frame for undercoverage, overcoverage, multiple coverage, clustering, and periodicity, and make adjustments where necessary. Ideally, the list will be in a random order with respect to the study variable or, better yet, ordered in terms of the variable of interest or its correlate, thereby creating implicit stratification. If the sampling frame is randomized, systematic sampling is considered to be a good approximation of simple random sampling.
5. Determine the number of elements in the sampling frame (N).
6. Calculate the sampling interval (i) by dividing the number of elements in the sampling frame (N) by the targeted sample size (n). One should ignore a remainder and round down or truncate to the nearest whole number. Rounding down and truncating may cause the sample size to be larger than desired. If so, one may randomly delete the extra selections. If the exact size of the population is not known and impractical to determine, one may fix the sampling fraction.
7. Randomly select a number, r, from "1" through i.
8. Select for the sample, r, $r + i$, $r + 2i$, $r + 3i$, and so forth, until the frame is exhausted.

At a technical level, systematic sampling does not create a truly random sample. It is often referred to as "pseudo random sampling," "pseudo simple random sampling," or "quasi-random sampling." Only the selection of the first element in systematic sampling is a probability selection. Once the first element is selected, some of the elements will have a zero probability of selection. Moreover, certain combinations of elements, such as elements that are adjacent to each other in the sampling frame, are not likely to be selected. Repeated systematic sampling, described below, may be used to address this problem.

What Are the Subtypes of Systematic Sampling?

Systematic sampling may be classified into three major types: linear systematic sampling, circular systematic sampling, and repeated (or replicated) systematic sampling. Linear systematic sampling is the most frequently used form of systematic sampling. The steps in selecting a linear systematic sample are those listed above. Circular systematic sampling may be viewed as a subtype of linear systematic sampling. In using this procedure, in Step 7, instead of selecting a random number between "1" and "*i*," the size of the interval, a random number is selected between "1" and "*N*." When one gets to the end of the list in selecting the sample, one would continue from the beginning of the list. This creates a circular pattern in selecting the sample.

Linear systematic sampling and circular systematic create a single sample. Repeated systematic sampling involves the selection of multiple samples from the target population and then combining them into a single sample. Instead of only one random start, several smaller systematic samples are selected using multiple random starts. This makes the process more time-consuming than linear systematic sampling. However, repeated sampling minimizes the effect of bias due to periodicity, a regularly occurring pattern in the sampling frame (see below). Moreover, because linear systematic sampling generates only one "cluster" of elements (although the cluster may contain multiple elements), technically, an unbiased estimate of sampling error cannot be obtained without making certain assumptions. At least two independently chosen clusters must be made. Repeated sampling provides more than one cluster of elements and facilitates the calculation of variances and standard error of estimates from the sample.

What Are the Strengths and Weaknesses of Systematic Sampling?

Systematic sampling has the strengths and weaknesses associated with most probability sampling procedures when compared to nonprobability sampling procedures. In highlighting the strengths and weaknesses of systematic sampling, we may compare it to simple random sampling. Systematic sampling is often used when it is impractical or impossible to use simple random sampling. When compared to simple random sampling, in some instances it is a stronger sampling procedure, and in other instances it is a weaker sampling procedure (see Table 5.8). Compared to simple random sampling, the strengths of systematic sampling include:

- If the selection process is manual, systematic sampling is easier, simpler, less time-consuming, and more economical than simple random sampling. One needs to use a random process to select only the first element. On the other hand, if the selection process is computerized, the ease in the selection process of systematic sampling and simple random sampling may be comparable to each other.
- If the sampling frame has a monotonic ordering that is related to a study variable (e.g., ordering of stores by dollar value, listing of employees by number of years employed, and listings of schools by graduation rates), implicit stratification may result in the statistical efficiency equivalent to that of proportionate stratified sampling and is thereby more efficient than simple random sampling. If the ordering is randomized, systematic sampling may yield results similar to simple random sampling.
- Systematic sampling ensures that the sample is more spread across the population.

Table 5.8 Strengths and Weaknesses of Systematic Sampling Compared to Simple Random Sampling

Strengths	Weaknesses
Unlike simple random sampling:	**Unlike simple random sampling:**
If the selection process is manual, systematic sampling is easier, simpler, less time-consuming, and more economical.	If the sampling interval is related to periodic ordering of the elements in the sampling frame, increased variability may result.
The target population need not be numbered and a sampling frame compiled if there is physical representation.	Combinations of elements have different probabilities of being selected.
If the ordering of the elements in the sampling frame is randomized, systematic sampling may yield results similar to simple random sampling.	Technically, only the selection of the first element is a probability selection since for subsequent selections, there will be elements of the target population that will have a zero chance of being selected.
If the ordering of the elements in the sampling frame is related to a study variable creating implicit stratification, systematic sampling is more efficient than simple random sampling.	Principle of independence is violated, for the selection of the first element determines the selection of all others.
Systematic sampling eliminates the possibility of autocorrelation.	Estimating variances is more complex than that for simple random sampling.
Systematic sampling ensures that the sample is spread across the population.	

Similarity of adjacent elements in a list makes for autocorrelation, the correlation among elements in the population. Although rare, this may occur in simple random sampling. Spatial autocorrelation is likely to exist in a listing of addresses. Persons who live at addresses that are close to each other are likely to be more similar to each other, say in terms of socioeconomic status, than they are to persons living at addresses that are not as close. A positive autocorrelation creates lower precision, and a negative autocorrelation creates higher precision when compared to simple random sampling. However, systematic sampling eliminates the possibility of autocorrelation. For example, in using a voter's list for the selection of a sample for a study of voter preferences, it is possible that members of the same family are selected using simple random sampling, but this is not possible using systematic sampling.

Compared to simple random sampling, systematic sampling has a number of weaknesses. Some of them include:

- Although its occurrence is relatively rare, periodicity in the sampling frame is a constant concern in systematic sampling. A biased sample could result if a periodic or cyclical pattern in the sampling frame corresponds to the sampling fraction. This problem will exist if the sampling fraction is equal to or a multiple of a periodic interval in the list. For example, a systematic sample of students would be biased if students are listed by class and within each class ranked by performance on an achievement test. If the classes have approximately the same number of students, periodic bias will result.
- Moreover, whereas in simple random sampling every combination of n elements has an equal chance of selection, this is not the case for systematic sampling.
- Technically only the selection of the first element is a probability selection since for subsequent selections there will be members of the target population that will have a zero chance of being selected.
- Principle of independence is violated, for the selection of the first element determines the selection of all the others.
- Estimating variances is more complex than that for simple random sampling.

Research Note 5.6 provides an example of systematic sampling. It describes the sampling procedures used by Chandek and Porter (1998) in their study of victims of robbery and burglary. Another example of systematic sampling is presented in Research Note 5.7. In this study, systematic sampling was used in selecting cases in a study of child abuse cases filed in Dallas, Texas, between December 2001 and December 2003.

RESEARCH NOTE 5.6

Example of Systematic Sampling: Study of Victims of Robbery and Burglary

Chandek and Porter (1998) utilized systematic sampling in their study of victims of robbery and burglary. They described their sampling procedures as follows:

> The data for this study were obtained from telephone surveys and official complainant records from a medium-size Midwestern police department. The sample was obtained from the total population of burglary and robbery victims whose crimes were reported to the department between May 15th and August 14th of 1995—a total of 2,000 burglary and 999 robbery victims. Systematic sampling procedures were used to create a manageable sample size given the project's resource constraints.
>
> After using systematic sampling procedures and eliminating cases with missing information on the official complainant records, cases where the victim was under the age of 18 and cases where the victim was a business rather than an individual, the sample comprised 216 robbery victims and 200 burglary victims. A telephone survey was then conducted using a questionnaire specifically designed for the present study.

Source: Chandek & Porter, 1998, pp. 26–27. Reprinted with permission.

RESEARCH NOTE 5.7

Example of Systematic Sampling: Study of Evidence and Filing of Charges in Child Abuse Cases

Walsh, Jones, Cross, and Lippert (2010) used systematic sampling in their study of the type of evidence and whether charges were filed in child abuse cases in Dallas, Texas. They described their sampling procedures as follows:

> Systematic sampling (e.g., taking every third case) was used to enroll research cases from the Children's Advocacy Center and from comparison community agencies (e.g., Child Protective Services, police). If there were multiple victims in the same family or multiple perpetrators per case, data collection focused

on one randomly selected victim or perpetrator. The initial sample for this analysis included only child sexual abuse cases with adult offenders (N = 360). Five cases were missing information on whether charges were filed; thus, they were not included. In 26 cases, an offender could not be identified; the offender or family fled during the investigation; or the family was unwilling to press charges. . . . The final sample included 329 cases.

Source: Walsh, Jones, Cross, & Lippert, 2010, pp. 440–441. Reprinted with permission.

CLUSTER SAMPLING

What Is Cluster Sampling?

Often it is impossible or impractical to create a sampling frame of a target population, and/or the target population is widely dispersed geographically, making data collection costs relatively high. Such situations are ideal for cluster sampling. **Cluster sampling** is a probability sampling procedure in which elements of the population are randomly selected in naturally occurring groupings (clusters). In the context of cluster sampling, a "cluster" is an aggregate or intact grouping of population elements. Element sampling is the selection of population elements individually, one at a time. On the other hand, cluster sampling involves the selection of population elements not individually, but in aggregates. The sampling units or clusters may be space-based, such as naturally occurring geographical or physical units (e.g., states, counties, census tracts, blocks, or buildings); organization-based, such as such units as school districts, schools, grade levels, or classes; or telephone-based, such as area codes or exchanges of telephone numbers. For the most part, the cluster sample designs described in this chapter are space-based or area-based sampling procedures. Telephone-based sampling procedures are described in Chapter 6.

The heterogeneity of the cluster is central to a good cluster sample design. Ideally, the within-cluster differences would be high, and the between-cluster differences would be low. The clusters should be like each other. On the other hand, the elements within each cluster should be as heterogeneous as the target population. Ideally, the clusters would be small but not so small as to be homogeneous.

What Are the Steps in Selecting a Cluster Sample?

There are six major steps in selecting a cluster sample:

1. Define the target population.
2. Determine the desired sample size.
3. Identify an existing sampling frame or develop a new sampling frame of clusters of the target population.
4. Evaluate the sampling frame for undercoverage, overcoverage, multiple coverage, and clustering, and make adjustments where necessary. Ideally, the clusters would be as heterogeneous as the population, mutually exclusive, and collectively exhaustive. Duplication of elements in the sample may result if population elements belonged to more than one cluster. Omissions will result in coverage bias.
5. Determine the number of clusters to be selected. This may be done by dividing the sample size by estimated average number of population elements in each cluster. To the extent the homogeneity and heterogeneity of the clusters are different from that of the population, as cluster number increases, precision increases. On the other hand, as differences between clusters increases, precision decreases.
6. Randomly select the targeted number of clusters.

What Are the Subtypes of Cluster Sampling?

Two major dimensions are used to classify different types of cluster sampling. One is based on the number of stages in the sample design, and the other is based on the proportional representation of the clusters in the total sample.

Subtypes Based on Number of Stages

Often cluster sampling is carried out in more than one "stage." A stage is a step in the sampling process in which a sample is taken. Considering the number of stages in the design, there are three major subtypes of cluster sampling: single-stage cluster sampling, two-stage cluster sampling, and multistage cluster sampling.

Single-stage cluster sampling. In a single-stage cluster sample design, sampling is done only once. As an example of single-stage cluster sampling, let us say one is interested in studying homeless persons who live in shelters. If there are five shelters in a city, a researcher will randomly select one of the shelters and then include in the study all the homeless persons who reside at the selected shelter. A market researcher might choose to use a single-stage cluster sample design. Say a researcher was interested in test marketing a product. The researcher may randomly select zip codes; send samples of the product together with a mail-back evaluation questionnaire to each address within the selected clusters.

Two-stage cluster sampling. A **two-stage cluster sample design** includes all the steps in single-stage cluster sample design with one exception, the last step. Instead of including all the elements in the selected clusters in the sample, a random sample (either a simple random sample, stratified sample, or systematic sample) is taken from the elements in each selected cluster. Sampling beyond the first stage is sometimes referred to as subsampling. Generally, unless the clusters are homogeneous, a two-stage cluster sample design is better than a one-stage cluster sample design. A self-weighting sample will result if at the first stage sampling is conducted with probability proportional to size (see below). Using the example of the study of homeless persons described above, instead of selecting all the persons who reside at the selected shelter for inclusion in the study, the researcher would randomly select a subset of the residents of the shelter.

Multistage cluster sampling. Surveys of large geographical areas require a somewhat more complicated sample design than those described up to this point. Typically, a multistage cluster sample design must be used. **Multistage cluster sampling** involves the repetition of two basic steps: listing and sampling. Typically, at each stage, the clusters get progressively smaller in size; and at the last stage element sampling is used. Sampling procedures (simple random sampling, stratified sampling, or systematic sampling) at each stage may differ. It is not necessary that the sampling procedures at each stage be the same. The number of stages that are used is often determined by the availability of sampling frames at different stages.

Special terminology is used to refer to the different sampling units. The sampling unit that is used in the first stage is referred to as the **primary sampling unit** (PSU). The units of subsequent sampling are referred to as the **secondary sampling unit** (SSU), tertiary sampling units (TSU), etc., until one gets to the "final" or "ultimate" sampling unit.

Typically, as the sampling process moves from the selection of PSUs to the other sampling stages, the sampling units become more homogeneous. The large clusters tend to be more heterogeneous than small clusters. Because of the greater heterogeneity of the PSUs, sampling error is minimized if one sample has more PSUs than SSUs, more SSUs than TSUs, and so forth.

Subtypes Based on the Proportional Representation of Clusters in Sample

Clusters may be selected in such a way that it is an EPSEM sampling procedure; that is, every element in the population would have an equal chance to be included in the sample. If the clusters sampled are roughly the same size, the sample design may be considered to be an EPSEM sample design. If the clusters have unequal sizes, an EPSEM sample design may be achieved by using a probability proportionate to size (PPS) selection procedure. The probability of selecting a cluster is dependent on the proportional distribution of its elements in the target population. Using PPS, a self-weighting sample is obtained. Probability disproportional to size (PDS) sampling involves selecting clusters without considering the proportional distribution of the elements in the target population.

Respondent Selection Procedures

Typically, in household surveys employing a two-stage cluster sample design or a multistage cluster sample design, individual elements are selected at the last stage of the sample design. If the household contains more than one member of the target population, one element must be selected. Both nonprobability and probability procedures are used to select the element from whom to collect data.

Two principal nonprobability household respondent selection procedures are used: head of household selection and first-adult selection. In using the head of household selection the researcher simply asks to speak to the head of household. One may alternatively ask for the male and female heads of household. The first-adult approach involves the selection of the first adult contacted, providing he/she is a member of the target population. These procedures are easy to administer, do not take much time, and are not intrusive. However, they incur selection bias, and are likely to oversample females as they are more likely than males to be available to be interviewed. The head of household method tends to oversample women, especially in urban areas, due to the greater number of single-parent female-headed households than single-parent male-headed households. The first-adult selection method tends to oversample women because women are more likely to be at home.

These respondent selection procedures do not give every member of the target population a chance to be included in the sample. Combining the probability selection of clusters with the nonprobability selection of household members makes the sampling procedure a mixed-methods procedure. Mixed-methods sampling procedures are described in more detail in the next chapter.

There are several probability household respondent selection procedures. The most frequently used probability approaches are the Kish tables, the Troldahl-Carter-Bryant tables, the Hagan and Carter selection method, and the last/next birthday method (Binson, Canchola, & Catania, 2000). These procedures reflect a struggle among researchers to minimize systematic error. Typically, the introduction to the interview is lengthened as they involve two consents: the initial consent from the first contact in the household and second from the person selected to be interviewed. This has the effect of decreasing undercoverage bias but increasing refusal rates. Moreover, if the selected person is not at home, the interviewer is restricted from selecting someone else in the household. Callbacks must be made. The success of the callbacks affects the study's unit nonresponse bias.

Kish Tables

In 1949, Kish created tables to facilitate the random selection of household members from among those eligible to participate in a study. The tables included a listing of household sizes one through five, and six or more; and for each household size a random number indicating the household member to be included in the study (see Table 5.9). Tables are prepared so that each household member (except those in households with six or more members of the target population) will have an equal chance to be selected and randomly applied to interview being conducted. Once making contact with a household, as part of the screening process, an interviewer would:

- Create a listing (sampling frame) of members of the household that are in the target population including their gender, relationship to household head, and age.
- Assign a unique number to each element listed in the frame.
- Using the randomized response table assigned to interview, determine the household member indicated in the table that should be interviewed.

Using the Kish tables produces a random sample to household members and decreases undercoverage bias; however, it does so at a cost. The process increases the amount of training of interviewers, the amount of time required

Table 5.9 Summary of Kish Tables Used for Selecting One Adult in Each Dwelling

Proportions of Assigned Tables	Table Number	If the number of adults in the household is:					
		1	2	3	4	5	6 or more
		Select adult numbered:					
1/6	A	1	1	1	1	1	1
1/12	B1	1	1	1	1	2	2
1/12	B2	1	1	1	2	2	2
1/6	C	1	1	2	2	3	3
1/6	D	1	2	2	3	4	4
1/12	E1	1	2	3	3	3	5
1/12	E2	1	2	3	4	5	5
1/6	F	1	2	3	4	5	6

Source: Kish, 1965, p. 399. Reprinted with permission.

for the screening process, the difficulty interviewers experience in establishing rapport, and the resistance of respondents to be interviewed. Due to the complexity of the method, some interviewers may improvise and use inappropriate shortcuts in selecting the person to be interviewed. The Kish tables were developed at a time when surveys were conducted primarily via personal interviews. As telephone surveys became more and more popular, the need for a less time-consuming respondent-selection procedure became more apparent. Other procedures were developed to satisfy such needs.

Troldahl-Carter-Bryant Tables

Troldahl-Carter-Bryant (TCB) tables are representative of a number of approaches designed to simplify the Kish tables (Bryant, 1975; Czaja, Blair, & Sebestik, 1982; Groves & Kahn, 1979; Paisley & Parker, 1965; Troldahl & Carter, 1964). Using TCB tables, a researcher asks only two questions: How many persons live in the household who are in the target population (say, 18 years of age or older), and how many of them are women? The TCB randomized response tables are then used by the interviewer in selecting to interview either the man, the woman, oldest man or woman, youngest man or woman, or the middle man or woman (for an example, see Table 5.10). As

Table 5.10 Example of Troldahl-Carter-Bryant Randomized Response Table

Number of Women in Household	Number of Adults in Household			
	1	2	3	4 or more
0	Man	Youngest man	Youngest man	Oldest man
1	Woman	Woman	Oldest man	Woman
2		Oldest woman	Man	Oldest man
3			Youngest woman	Man or oldest man
4 or more				Oldest woman

done for the Kish tables, these tables are randomly assigned to the interviews conducted. The Kish tables and the TCB tables provide a means of randomly selecting the person to be interviewed; however, each member of the target population does not have an equal chance to be selected. Although believed to be a minor violation of randomness, the TCB method does not allow the selection of persons who fall between the youngest and oldest persons.

Hagan and Carter Selection Method

Hagan and Collier (1983) used an even simpler method. Their approach involves the random assignment of four forms to the interviews that are conducted. One form instructs the interviewer to ask to speak with the youngest adult male, another instructs the interviewer to ask to speak to the youngest adult female, another instructs the interviewer to speak to the oldest adult male, and the fourth instructs the interviewer to speak to the oldest adult female. If no such person is present, the interviewer asks to speak to the opposite sex of the same age group. In order to compensate for the greater difficulty in contacting men and younger females, these subpopulations are often given higher probabilities to be selected. This procedure is easier than using the Kish tables and the TCB tables. However, it assumes that only two members of the target population are in the household and, as for TCB tables, does not allow the selection of persons who are inbetween the oldest and youngest persons in the household.

Last/Next Birthday Method

The approach developed by Salmon and Nichols (1983) does not involve enumeration of household members nor randomized selection tables. A researcher merely asks to speak to the member of the target population who had the last birthday or will have the next birthday. One may randomly alternate either the last or next birthday. Compared to the procedures described above, this method is the easiest and the least time-consuming in terms of the training required and its administration. However, the validity of the procedure is dependent on whether the person answering the screening questions actually knows the birthday of all the members of the household. The larger the household, the more likely this the person does not know the birthday of all eligible persons for the study. Moreover, the procedure is considered a quasi-probability procedure because the respondent is determined when the date to conduct the interview is determined.

Alphabetic Ordering of Names

Another approach is the alphabetic ordering of the first names of those in the target population. As the last, or next, birthday method, this method is relatively easy to administer but is dependent on the knowledge of the contact person. Moreover, it requires time to obtain the names, put them in alphabetical order, and then to make a selection.

What Are the Strengths and Weaknesses of Cluster Sampling?

Cluster sampling has the strengths and weaknesses associated with most probability sampling procedures when compared to nonprobability sampling procedures. However, it has several special strengths and weaknesses when compared to other probability sampling procedures, such as simple random sampling (see Table 5.11). Some of the strengths of cluster sampling when compared to simple random sampling are:

- If the clusters are geographically defined, cluster sampling requires less time, money, and labor than simple random sampling. It is the most cost-effective probability sampling procedure.
- For the same level of costs, cluster sampling with a higher sample size may yield less sampling error than that resulting from simple random sampling with a smaller sample size.

Table 5.11 Strengths and Weaknesses of Cluster Sampling Compared to Simple Random Sampling

Strengths	Weaknesses
Compared to simple random sampling:	**Compared to simple random sampling:**
If the clusters are geographically defined, cluster sampling requires less time, money, and labor.	A cluster sample may not be as representative of the population as a simple random sample of the same sample size.
Cluster sampling permits subsequent sampling because the sampled clusters are aggregates of elements.	Variances of cluster samples tend to be much higher than variances of simple random samples.
One can estimate characteristics of the clusters as well as the population.	Cluster sampling introduces more complexity in analyzing data and interpreting results of the analyses.
Cluster sampling does not require a sampling frame of all of the elements in the target population.	Cluster sampling yields larger sampling errors for samples of comparable size than other probability samples.

- Cluster sampling permits subsequent sampling because the sampled clusters are aggregates of elements.
- Unlike simple random sampling, cluster sampling permits the estimation characteristics of subsets (clusters) as well as the target population.
- Single-stage cluster sampling requires a sampling frame of the clusters only, and two-stage cluster sampling and multistage cluster sampling require a sampling frame of the elements of the population only for the clusters sampled at the last stage of the process.
- Cluster sampling is much easier to implement than simple random sampling.

Some of the weaknesses of cluster sampling when compared to simple random sampling include:

- The sampled clusters may not be as representative of the population as a simple random sample of the same sample size.
- Combining the variance from two separately homogeneous clusters may cause the variance of the entire sample to be higher than that of simple random sampling.
- Cluster sampling introduces more complexity in analyzing data. Inferential statistical analysis of data collected via cluster sampling is more difficult to

compute and interpret results than inferential statistical analysis of data collected via simple random sampling. The statistical software used to analyze the data collected must use formulas that take into account the use of a cluster sample design. Many statistical software programs utilize formulas for simple random sampling and, as a result, overestimate levels of significance.

- The more stages there are in a cluster sample design, the greater overall sampling error.
- If clusters are not similar to each other, the fewer the number of clusters, the greater the sampling error.
- Cluster sampling yields larger sampling errors for samples of comparable size than other probability samples. If the clusters are similar to each other, this error is minimized. Moreover, these errors can be reduced by increasing the number of clusters. Note, this has the effect of increasing data collection costs.
- The more clusters one selects, the less the difference in data collection costs between cluster sampling and simple random sampling.
- Since elements within a cluster tend to be alike, we receive less new information about the population when we select another element from that cluster rather than from another cluster. This lack of new information makes a cluster sample less precise than a simple random sample.

What Is the Difference Between Cluster Sampling and Stratified Sampling?

Cluster sampling is similar to stratified sampling in that both involve separating the population into categories and then sampling within the categories (see Table 5.12). Both sampling procedures permit analysis of individual categories (strata or clusters) in addition to analysis of the total sample. However, there are important differences. Some of these differences include:

- In stratified sampling, once the categories (strata) are created, a random sample is drawn from each category (stratum). On the other hand, in cluster sampling, elements are not selected from each cluster. In single-stage cluster sampling, once the categories (clusters) are created, a random sample of cluster is drawn. All elements in the selected cluster are included in the sample. In two-stage cluster sampling and multi-stage cluster sampling, a random sample of cluster is drawn and then elements are randomly selected from the selected clusters.

Table 5.12 Comparison of Stratified Sampling and Cluster Sampling

Stratified Sampling	**Cluster Sampling**
The population is separated into strata, and then sampling is conducted within each stratum.	The population is separated into clusters, and then clusters are sampled.
Analysis of individual strata is permitted in addition to analysis of the total sample.	Analysis of individual categories (clusters) are permitted in addition to analysis of the total sample.
In order to minimize sampling error, within-group differences among strata should be minimized, and between/group differences among strata should be maximized.	In order to minimize sampling error, within-group differences should be consistent with those in the population, and between-group differences among the clusters should be minimized.
A sampling frame is needed for the entire target population.	In single-state cluster sampling, a sampling frame is needed only for the clusters. In two-stage and multistage cluster sampling, a sampling frame of individual elements is needed only for the elements in the clusters selected at the final stage.
Main purpose: increase precision and representation.	Main purpose: decrease costs and increase operational efficiency.
Categories are imposed by the researcher.	Categories are naturally occurring pre-existing groups.
More precision compared to simple random sampling.	Lower precision compared to simple random sampling.
The variables used for stratification should be related to the research problem.	The variables used for clustering should not be related to the research problem.
Common stratification variables: age, gender, income, race.	Common classification variables: geographical area, school, grade level.
Requires more prior information than cluster sampling.	Requires less prior information than stratified sampling.

- In stratified sampling, in order to minimize sampling error, within-group differences among strata should be minimized, and the strata should be as homogeneous as possible. In cluster sampling, in order to minimize sampling error, within-group differences should be consistent with those in the population, and the clusters should be as heterogeneous as the population. The ideal situation for stratified sampling is to have the homogeneity within each stratum and the strata means to differ from each other. The ideal situation for cluster sampling is to have heterogeneity within the clusters and the cluster means not to differ from each other.
- In stratified sampling, in order to minimize sampling error, between-group differences among strata should be maximized. In cluster sampling, in order to minimize sampling error, between-group differences among the clusters should be minimized.
- In stratified sampling, categories are conceptualized by the researcher. In cluster sampling, the categories are naturally occurring groups.
- In stratified sampling, a sampling frame is needed for the entire target population. In single-stage cluster sampling, a sampling frame is needed only for the clusters. In two-stage cluster sampling and multistage cluster sampling, in addition to a sampling frame of the clusters in the first stage of the process, a sampling frame is needed only for elements of each one of the selected clusters.
- The main purpose of stratified sampling is to increase precision and representativeness. The main purpose of cluster sampling is to decrease costs and increase operational efficiency.
- Compared to simple random sampling, stratified sampling has higher precision and cluster sampling has lower precision. The increase in precision by stratification is not that much. However, clustering can cause a significant decrease in precision.
- The variables used for stratification should be related to the variables under study. The variable used for clustering should not be related to the variables under study.
- Commonly used stratification variables are age, gender, and income. Commonly used classification variables in cluster sampling are geographical area, school, and grade level.
- Stratified sampling requires more prior information than cluster sampling; likewise, cluster sampling requires less prior information than stratified sampling.
- In stratified sampling, the researcher strives to divide the target population into a few subgroups, each with many elements in it. In cluster sampling, the researcher strives to divide the target population into many subgroups, each with few elements in it.

What Is the Difference Between Multistage Sampling and Multiphase Sampling?

Multistage sampling (two-stage cluster sampling and multistage cluster sampling) is often confused with multiphase sampling (also referred to as two-phase sampling, double sampling, and post-stratification sampling). Both sampling procedures involve the multiple sampling at different stages or phases, and in some circumstances may be viewed as mixed-methods sampling. In multistage sampling the sampling units for the different stages are different. On the other hand, in multiphase sampling the same sampling unit is sampled multiple times.

Typically, multiphase sampling is used when one does not have a sampling frame with sufficient auxiliary information to allow for stratification. The first phase is used for screening purposes. Using the available sampling frame, one may proceed as follows:

1. Select an initial sample of elements from the available sampling frame.
2. Conduct a short screening interview to collect the necessary auxiliary information for further sampling and stratification.
3. Poststratify the initial sample into strata using the auxiliary information collected.
4. Using the strata for which one desires to collect additional information, select either all the elements in the strata or a probability sample of the elements in the strata for additional data collection.

Multiphase sampling typically is carried out to increase precision, reduce costs, and reduce nonresponse. As noted earlier, stratified samples have higher levels of precision than simple random samples of the same sample size. However, a sampling frame must include information on the stratification variable(s) for all population elements to employ stratification. Multiphase sampling is an option when a sampling frame does not include such information.

Multiphase sampling may also be employed to reduce data collection costs if it took more time and effort to collect data on some variables than to collect data on other variables. In Phase 1, the easily accessible data may be collected from the entire sample. In Phase 2 and other subsequent phases, if desired or necessary, the data that take greater effort or expense to be collected are collected from a smaller subsample. Data collection costs are minimized.

Multiphase sampling may also be used to obtain information on nonrespondents.

Typically, it costs more to collect data on persons who initially refused to participate in a study and other nonrespondents than to collect data from the initial respondents. Such costs might be minimized by employing a multiphase sampling of nonrespondents.

Below are descriptions of several popular national surveys that are representative of multistage cluster sampling: the National Home and Hospice Care Survey (Research Note 5.8), the National Ambulatory Medical Care Survey (Research Note 5.9), the National Health and Nutrition Examination Survey (Research Note 5.10), the National Survey of Family Growth (Research Note 5.11), and the National Health Interview Survey (Research Note 5.12).

RESEARCH NOTE 5.8

Example of Two-Stage Cluster Sampling With Probability Proportional to Size: The National Home and Hospice Care Survey

The National Home and Hospice Care Survey (NHHCS) is a continuing series of surveys of home and hospice care agencies in the United States. Data have been collected about agencies that provide home and hospice care and about their current patients and discharges. Beginning in 1992, the survey was repeated in 1993, 1994, 1996, 1998, and 2000, and most recently in 2007. The 2007 NHHCS used a stratified two-stage probability sample design with probability proportional to size. The sample design included these two stages:

> The first stage, carried out by the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS), was the selection of home health and hospice agencies from the sample frame of over 15,000 agencies, representing the universe of agencies providing home health care and hospice services in the United States. The primary sampling strata of agencies were defined by agency type and metropolitan statistical area (MSA) status. Within these sampling strata, agencies were sorted by census region, ownership, certification status, state, county, ZIP code, and size (number of employees). For the 2007 NHHCS, 1,545 agencies were systematically and randomly sampled with probability proportional to size. . . . The second stage of sample selection was completed by the interviewers during the

agency interviews. The current home health patients and hospice discharges were randomly selected by a computer algorithm, based on a census list provided by each agency director or his/her designee. Up to 10 current home health patients were randomly selected per home health agency, up to 10 hospice discharges were randomly selected per hospice agency, and a combination of up to 10 current home health patients and hospice discharges were randomly selected per mixed agency.

Source: Centers for Disease Control and Prevention, 2007.

RESEARCH NOTE 5.9

Example of Multistage Cluster Sampling: The National Ambulatory Medical Care Survey

The National Ambulatory Medical Care Survey (NAMCS) is a national survey designed to collect data on the provision and use of ambulatory medical care services in the United States. The survey involves the sampling of visits to non-federally employed office-based physicians who are primarily engaged in direct patient care. The survey was conducted annually from 1973 to 1981, in 1985, and annually since 1989. It utilizes the following multistage sample design:

> The NAMCS utilizes a multistage probability design that involves probability samples of **primary sampling units** (PSUs), physician practices within PSUs, and patient visits within practices. The first-stage sample includes 112 PSUs. PSUs are geographic segments composed of counties, groups of counties, county equivalents (such as parishes or independent cities) or towns and townships (for some PSUs in New England) within the 50 States and the District of Columbia.
>
> The second stage consists of a probability sample of practicing physicians selected from the master files maintained by the American Medical Association and the American Osteopathic Association. Within each PSU, all eligible physicians were stratified by 15 groups: general and family practice, osteopathy, internal medicine, pediatrics, general surgery, obstetrics and gynecology, orthopedic surgery, cardiovascular diseases, dermatology, urology, psychiatry, neurology, ophthalmology, otolaryngology, and a residual category of all other specialties.

(Continued)

(Continued)

The final stage is the selection of patient visits within the annual practices of sample physicians. This involves two steps. First, the total physician sample is divided into 52 random subsamples of approximately equal size, and each subsample is randomly assigned to 1 of the 52 weeks in the survey year. Second, a systematic random sample of visits is selected by the physician during the reporting week. The sampling rate varies for this final step from a 100 percent sample for very small practices, to a 20 percent sample for very large practices as determined in a presurvey interview.

Source: Centers for Disease Control and Prevention, 2010.

RESEARCH NOTE 5.10

Example of Multistage Cluster Sampling: The National Health and Nutrition Examination Survey (NHANES)

The National Health and Nutrition Examination Survey (NHANES) is a series of studies designed to assess the health and nutritional status of adults and children in the United States. The series began in the early 1960s. Its sample design consists of the following stages:

Stage 1: Primary sampling units (PSUs) are selected. These are mostly single counties or, in a few cases, groups of contiguous counties with probability proportional to a measure of size (PPS).

Stage 2: The PSUs are divided up into segments (generally city blocks or their equivalent). As with each PSU, sample segments are selected with PPS.

Stage 3: Households within each segment are listed, and a sample is randomly drawn. In geographic areas where the proportion of age, ethnic, or income groups selected for oversampling is high, the probability of selection for those groups is greater than in other areas.

Stage 4: Individuals are chosen to participate in NHANES from a list of all persons residing in selected households. Individuals are drawn at random within designated age-sex-race/ethnicity screening subdomains. On average, 1.6 persons are selected per household.

Source: Centers for Disease Control and Prevention, 2010.

RESEARCH NOTE 5.11

Example of Multistage Cluster Sampling: The National Survey of Family Growth

Beginning in 1971, the National Survey of Family Growth (NSFG) obtains detailed information on factors affecting childbearing, marriage, and parenthood from a national probability sample of women and men 15 to 44 years of age. Its 2006 to 2010 sample design consists of the following steps:

> The NSFG sample design consisted of *five stages* of selection to choose eligible sample persons. Women, teens 15–19 years of age, and black and Hispanic persons are selected at higher rates, yielding an oversample of such persons.
>
> The 2006–2010 NSFG sample design started with the same national sample of PSUs used in the 2002 ("Cycle 6") NSFG national sample design. . . . Following the creation of the PSUs, a process called stratification was used to partition the PSUs into three major groups or strata: 28 large metropolitan areas, 290 other metropolitan areas, and 2,084 nonmetropolitan areas. The 28 large metropolitan areas are referred to as self-representing (SR) areas. SR areas are those that have such large populations that a national sample of the size used for continuous NSFG virtually required that they be represented. As such "certainty" selections, the sample from each of these areas represents only those areas. That is, the sample from these 28 PSUs represents only the population of that area. Hence, in the sampling literature, these types of units are referred to as representing only themselves, or "self-representing."
>
> The remaining 2,374 PSUs are called non-self-representing (NSR) areas. A sample of the NSR PSUs was selected so that each sample PSU represented itself and other NSR PSUs of a similar nature. In order to make the representation more complete, the NSR PSUs were further grouped by geography and population size into 82 sets or strata. Each NSR stratum had two or more PSUs, and some strata had more than 100 PSUs. The number of PSUs in a stratum varied because the strata were created to have approximately equal 2000 Census population across the PSUs. . . In the second stage of selection, census blocks were stratified into four domains within each PSU, and the housing units on those blocks were listed. . . . The third stage of selection chooses housing units from the list of addresses available

(Continued)

(Continued)

> in each sample segment. . . . The fourth stage of sampling is the selection of eligible persons within sample households. Interviewers visit housing units selected in the third stage, and when the housing unit is found to be occupied, attempt to list all persons living there. One eligible person is chosen randomly in every household containing one or more eligible persons. . . . The fifth stage in sample selection occurs in each 12-week "quarter" of interviewing: the selection of the "double sample" (because it is a sample of a sample). After 10 of the 12 weeks of data collection in each 12-week quarter, a set of selected housing units has not been successfully screened or, if successfully screened, the sampled person has not been interviewed yet.

Source: Lepkowski, Mosher, Davis, Groves, & Van Hoewyk, 2010, pp. 5–6. Reprinted with permission.

RESEARCH NOTE 5.12

Example of Multistage Cluster Sampling: Sample Design of the National Health Interview Survey

Since 1957, the National Health Interview Survey has been the principal source of health information on the U.S. population. In 2006, it utilizes a multistage area probability design:

> The National Health Interview Survey is a cross-sectional household interview survey. Sampling and interviewing are continuous throughout each year. The sampling plan follows a multistage area probability design that permits the representative sampling of households and noninstitutional group quarters (e.g., college dormitories) . . . The first stage of the current sampling plan consists of a sample of 428 primary sampling units (PSUs) drawn from approximately 1,900 geographically defined PSUs that cover the 50 States and the District of Columbia. A PSU consists of a county, a small group of contiguous counties, or a metropolitan statistical area.
>
> Within a PSU, two types of second-stage units are used: area segments and permit segments. Area segments are defined geographically and contain an expected eight, twelve, or sixteen addresses. Permit segments cover

housing units built after the 2000 census. The permit segments are defined using updated lists of building permits issued in the PSU since 2000 and contain an expected four addresses. . . . As with the previous sample design, the NHIS sample is drawn from each State and the District of Columbia. Although the NHIS sample is too small to provide State level data with acceptable precision for each State, selected estimates for most states may be obtained by combining data years.

The total NHIS sample is subdivided into four separate panels, or subdesigns, such that each panel is a representative sample of the U.S. population. This design feature has a number of advantages, including flexibility for the total sample size. For example, the 2006 and 2007 NHIS samples both were reduced because of budget shortfalls; two panels were cut from the sample in the third calendar quarter of each year.

The households and noninstitutional group quarters selected for interview each week in the NHIS are a probability sample representative of the target population. With four sample panels and no sample cuts, the expected NHIS sample size (completed interviews) is approximately 35,000 households containing about 87,500 persons.

Source: Centers for Disease Control and Prevention, 2009.

GUIDELINES FOR CHOOSING TYPE OF PROBABILITY SAMPLE DESIGN

Considering the features of the sample designs described above, and their strengths and weaknesses, the following guidelines may be offered for using the following types of sample designs:

- Simple random sampling
- Stratified sampling
- Systematic sampling
- Cluster sampling
- Mixed sample designs

Guideline 5.1. *Simple random sampling.* Consider choosing simple random sampling if one has access to a complete and accurate sampling frame of the target population that is complete and accurate but does not contain auxiliary information that may be used for stratification purposes.

Guideline 5.2. *Stratified sampling*. Consider choosing stratified sampling if:

- It is possible to divide a population into two or more homogeneous strata and construct a sampling frame for each stratum.
- One has access to a sampling frame of the target population that is complete and accurate and contains auxiliary information that may be used for stratification purposes.
- Some subgroups of the population are vastly different from other subgroups.
- It is very important to minimize sampling error.
- There is a concern about underrepresenting smaller subgroups.
- The population is heterogeneous.
- There is a desire to use different selection methods for different strata.
- It is likely that answers to the research questions of a study are likely to be different for different subgroups.
- It is useful when each stratum needs to be reported separately.
- Comparative analysis of strata is desired.

Guideline 5.3. *Proportionate stratified sampling.* Consider choosing proportionate stratified sampling if subgoups of approximately the same size are to be investigated or compared.

Guideline 5.4. *Disproportionate stratified sampling.* Consider choosing disproportionate stratified sampling if:

- Subgoups of vastly different sizes are to be investigated or compared.
- It is important to include a large number of elements from a small segment of the population.
- One is primarily interested in key similarities and differences among strata.
- Some observations are limited or hard to obtain.
- It is important to make statistically valid statements about subgroups.
- Subgroups of the population have different variances for the variables of interest.
- Costs of data collection are different across population subgroups.

Guideline 5.5. *Systematic sampling.* Consider choosing systematic sampling if:

- It is difficult to identify items using a simple random sampling method.
- It is important to use a probability sampling procedure that can be easily implemented.
- A sampling frame is not available or impractical to prepare, but a stream of representative elements of the population is available.
- The listing of the population is essentially random or can be randomized.

Guideline 5.6. *Cluster sampling.* Consider choosing cluster sampling if:

- It is important to minimize data collection costs and there are substantial fixed costs associated with each data collection location.
- A sampling frame of individual population elements is not available but a sampling frame of clusters of elements is available.
- Travel costs can be substantially reduced.

Guideline 5.7. *Double sampling and multiphase sampling.* Consider double sampling and multiphase sampling if there is a need to identify and collect information from a subgroup of the population that is difficult to collect prior information on.

SUMMARY

There are four major choices of probability sample designs: simple random sampling, stratified sampling, systematic sampling, and cluster sampling. The strengths and weaknesses of the above sample designs are compared, and guidelines are presented for their selection.

Simple random sampling is a probability sampling procedure that gives every element in the target population and each possible sample of a given size, an equal chance of being selected. As with other probability sampling procedures, it tends to yield representative samples, and allows the use of inferential statistics to compute margin of errors. However, it tends to have larger sampling errors and less precision than stratified samples of the same sample size. If the target population is widely dispersed, data collection costs might be higher for simple random sampling than those for other probability sample designs, such as cluster sampling.

Stratified sampling is a probability sampling procedure in which the target population is first separated into mutually exclusive, homogeneous segments (strata), and then a simple random sample is selected from each segment (stratum). There are two major subtypes of stratified sampling: proportionate stratified sampling and disproportionate stratified sampling. In proportionate stratified sampling, the number of elements allocated to the various strata is proportional to the representation of the strata in the target population. This condition is not satisfied in disproportionate stratified sampling. In this type of stratification, unequal disproportionate allocation, equal disproportionate allocation, or optimum allocation may be applied.

Compared to unstratified sampling, stratified sampling (1) permits the estimation of population parameters and within-strata inferences and comparisons across strata; (2) tends to be more representative of a population; (3) takes advantage of knowledge the researcher has about the population; (4) possibly makes for lower data collection costs; and (5) permits the researcher to use different sampling procedures within the different strata. On the other hand, unlike unstratified sampling, stratified sampling requires prior information on the stratification variables and more complex analysis procedures.

Systematic sampling is a probability sampling procedure in which a random selection is made of the first element for the sample, and then subsequent elements are selected using a fixed or systematic interval until the desired sample size is reached. Generally, systematic sampling is easier, simpler, less time-consuming, and more economical than simple random sampling. If the ordering is unrelated to the study variables, but randomized, systematic sampling will yield results similar to simple random sampling. On the other hand, periodicity in the sampling frame is a constant concern in systematic sampling.

Cluster sampling is a probability sampling procedure in which elements of the population are randomly selected in naturally occurring aggregates or clusters. Subtypes of cluster sampling may be classified on the basis of the number of sampling events (single-stage cluster sampling, two-stage cluster sampling, and multistage cluster sampling) and on the basis of the proportional representation of the clusters in the sample (probability proportional to size and probability disproportional to size). Some of the strengths of cluster sampling when compared to simple random sampling include requiring less time, money, and labor; and permitting subsequent sampling and the estimation characteristics of clusters as well as the target population. However, cluster sampling when compared to simple random sampling may not be as representative of the population as a simple random sample of the same sample size, and variances of cluster sampling are likely to be higher than those for simple random sampling.

REVIEW QUESTIONS

1. What are the principal differences and similarities between the major categories of probability sampling: simple random sampling, systematic sampling, stratified sampling, and cluster sampling?
2. What are the principal differences and similarities among the subcategories of these major categories of probability sampling?
3. Provide examples of research questions for which the different types of stratified sampling would be a good fit.
4. Which strengths and weaknesses of the major types of probability sampling are the most critical?
5. What are the similarities and differences between stratified sampling and quota sampling?
6. What guidelines should be followed in establishing the strata for a stratified sample?

7. Compare and contrast procedures that have been used to select a respondent to be interviewed from those in a household who are eligible to participate in a study.
8. What is a sampling frame? Is it necessary to use a sampling frame in selecting a probability sample? Justify your answer.
9. What are the similarities and differences between cluster sampling and stratified sampling?
10. What are the similarities and differences between multistage sampling and multiphase sampling?
11. What do you consider to be the most critical guidelines for choosing among the various type of probability sampling? Why?
12. What guidelines should be considered in deciding between:
    a. Simple random sampling versus stratified sampling
    b. Element sampling versus cluster sampling
    c. Simple random sampling versus systematic sampling
13. A stratified sample of size n = 60 is to be taken from a population of size N = 4000, which consists of three strata of size N_1 = 2000, N_2 = 1200 and N_3 = 800. If the allocation is to be proportional, how large a sample must be taken from each stratum?
14. What alternative sample designs would you propose for the sample designs described in the research notes in this chapter?
15. Consider the sample designs described in the research notes in Chapter 4: Choosing the Type of Nonprobability Sample Design. What alternative probability sample designs may be used to achieve the purposes of the study? Compare and contrast the advantages and limitations of the probability sample designs you propose with the nonprobability sample designs described in the research notes in Chapter 4.
16. What procedures would you use to select a probability sample of homeless people and why? Once you have answered these questions, consider Burnam and Koegel's "Methodology for Obtaining a Representative Sample of Homeless Persons: The Los Angeles Skid Row Study" (1988).

KEY TERMS

Define and give examples of the following concepts:

cluster sampling

disproportionate allocation

equal probability selection method

multistage cluster sampling

optimum allocation
primary sampling unit
proportionate allocation
secondary sampling unit
simple random sampling
single-stage cluster sampling
systematic sampling
two-stage cluster sampling

REFERENCES FOR FURTHER STUDY

Binson, D., Canchola, J. A., & Catania, J. A. (2000). Random selection in a national telephone survey: A comparison of the Kish, next birthday, and last-birthday methods. *Journal of Official Statistics, 16, 53–59*.

Burnam, M. A., & Koegel, P. (1988). Methodology for obtaining a representative sample of homeless persons: The Los Angeles Skid Row Study. *Evaluation Review, 12*, 117–152.

Bryant, B. E. (1975). Respondent selection in a time of changing household composition. *Journal of Marketing Research, 12,* 129–135.

Czaja, R., Blair, J., & Sebestik, J. P. (1982). Respondent selection in a telephone survey: A comparison of three techniques. *Journal of Marketing Research, 19,* 381–385.

Hagan, D. E., & Collier, C. M. (1983). Must respondent selection procedures for telephone surveys be invasive? *Public Opinion Quarterly, 47*, 547–556.

Kish, L. (1949). A procedure for objective respondent selection within the household. *Journal of the American Statistical Association, 44,* 380–387.

Kish, L. (1965). *Survey sampling*. New York: Wiley & Sons.

Lavrakas, P. J., Bauman, S. L., & Merkle, D. M. (1993). The last-birthday method and within-unit coverage problems. *Proceedings of the section on survey research methods, American Statistical Association*, 1107–1112.

Levy, P. S., & Lemeshow, S. (2008). *Sampling of populations: Methods and applications*. New York: Wiley & Sons.

Salmon, C. T., & Nichols, J. S. (1983). The next-birthday method of respondent selection. *Public Opinion Quarterly, 47*, 270–276.

Scheaffer, R. L., Mendenhall, W., & Ott, L. (2006). *Elementary survey sampling*. Belmont, CA: Duxbury Press.

Sudman, S. (1976). *Applied sampling*. New York: Academic Press.

Thompson, S. K. (2002). *Sampling*. New York: Wiley & Sons.

Troldahl, V. C., & Carter, R. E. (1964). Random selection of respondents within households in phone surveys. *Journal of Marketing Surveys, 1,* 71–76.