2

Types of Error in Surveys

Surveys are designed to produce statistics about a target population. The process by which this is done rests on inferring the characteristics of the target population from the answers provided by a sample of respondents. This chapter outlines the two kinds of inferences that are required. It also describes the two types of error, bias and variability, that can limit the accuracy of those inferences.

Two of the main goals of survey methodology are to minimize error in data collected by surveys and to measure the error that necessarily is part of any survey. Chapters 3 through 9 describe in detail strategies for minimizing and measuring error. In order to appreciate those chapters, and to understand survey methodology, it is first necessary to understand what we mean by *error.*

As we have said, the purpose of a survey is to provide statistical estimates of the characteristics of a target population, some set of people. To do that we designate a subset of that population, a sample, from whom we try to collect information. *One fundamental premise of the survey process is that by describing the sample of people who actually respond, one can describe the target population.* The hope is that the characteristics the survey is designed to describe are present to the same degree, and are distributed in the same way, in the sample responding as in the target population as a whole.

The other defining characteristic of a survey is that respondents answer questions. The answers to the questions are used to describe the experiences, opinions, and other characteristics of those answering the questions. *A second fundamental premise of the survey research process is that the answers people give can be used to accurately describe characteristics of the respondents.* The extent to which those answers are not accurate measures is the second fundamental source of error in surveys.

Figure 2.1 shows in graphic form the way analysis of survey data works and the inferences on which it is based. The goal is to learn the characteristics of the target population. The material we have to work with consists of the answers the respondents gave in the survey. We tabulate the answers and would like to make the assumption that the answers are accurate measures of the characteristics of the respondents we are trying to measure. We then would

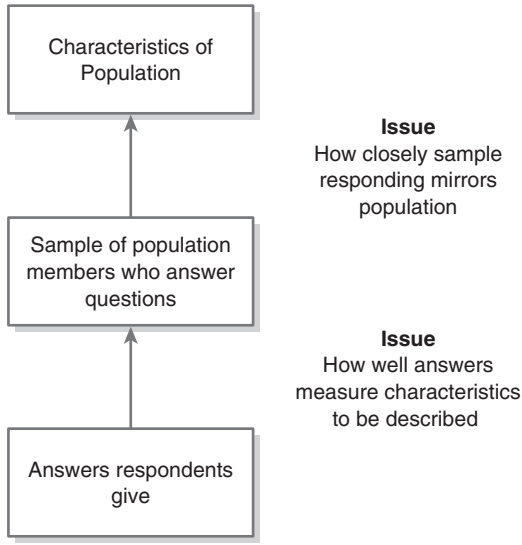

Figure 2.1 Inference in Survey Research

like to be able to further assume that by describing the sample of respondents, we are accurately describing the entire target population.

Some aspects of survey methodology are designed to address how closely a sample of respondents mirrors the population. Some aspects of survey methodology are designed to address how well the answers to the questions collected in the survey serve as measures of what they are intended to measure. The design of the survey and the way data collection is carried out can affect one or both of these potential sources of error.

ERROR ASSOCIATED WITH WHO ANSWERS

Any time a sample is drawn from a larger population, there is some chance that by chance alone that sample will differ from the total population from which it was drawn. A simple example that statisticians like to use is flipping a coin, which is heads on one side and tails on the other. Even if the coin is

perfectly fair, a sample of 10 flips will not always produce 5 heads and 5 tails. While 5 heads and 5 tails will be most common, a certain number of samples of 10 flips will produce 6 heads or 4 heads; 3 and 7 heads will be less common; 8 or 2 heads will be even less common; even more extreme distributions will be increasingly less common, but even those will occur if enough samples of 10 flips of a coin are actually tried. In the same way, if a population consists of 50% males and 50% females, any particular sample may by chance have more or fewer females than one would expect from the population as a whole.

In a sample survey, we usually only have a single sample from which to generalize. By chance, that sample can and will differ slightly from what it would look like if it perfectly mirrored the distribution of characteristics in the population. One goal of survey methodology is to minimize the random differences between the sample and the population. The way the sample is designed and selected can affect how closely the sample is likely to mirror the characteristics of the population from which it is drawn.

One kind of error of concern to survey methodologists is this random variation from the true characteristics of the population. This variation, the possible error that stems solely from the fact that data are collected from a sample rather than from every single member of the population, is called *sampling error.* It is one kind of error survey methodologists try to minimize.

A second kind of error that affects the relationship between a sample of respondents and of the population is *bias.* Bias means that in some systematic way the people responding to a survey are different from the target population as a whole.

There are three steps in the process of collecting data from a sample, each of which could, potentially, introduce bias into a sample:

1. The first step involves choosing the *sample frame,* those who actually have a chance to be selected. If there are some people in the target population who do not have any chance at all to be selected for the sample, and if they are somehow consistently different from those who do have a chance to be selected, the resulting sample will be biased in those ways. As an example, most surveys in the United States leave out people who live in group homes, such as prisons, convents, and nursing homes, and they leave out people who have no home address at all. Most telephone surveys not only leave out those without any telephone service at all but also those households that only have cell phones. For variables on which those people who are included are different from those who are systematically left out, the samples from whom data are collected will be biased as well.

2. If somehow the *process of selecting* who is in the sample is not random, the result could be a sample of respondents who are different from the target population as a whole. For example, if a sample consists of people who volunteer to be in a survey, they are likely to have a different profile of interests than those who do not volunteer.

3. Finally, *failure to collect answers from everyone* selected to be in the sample is a third potential source of bias. Some people are not available to answer questions; some people are unable to answer questions, due to their health or their language skills; some people refuse to answer the questions. To the extent that those who are unavailable, unable, or unwilling to answer questions are different from the rest of the population in ways that affect the survey answers, the results of the survey may be biased.

It is important to understand the distinction between the two kinds of errors in data. Sampling error, the first kind of error that was discussed, is random error. By chance, sometimes there will be too many females in the sample, sometimes too few, but on average, a series of properly drawn samples will have very close to the same percentage of females as the population as a whole. A challenge for methodologists is to minimize the variability from sample to sample to increase the likelihood that any given sample is very close to the population as a whole.

In contrast, features of the design or execution of a survey that bias samples will, on average, produce estimates that are consistently different from the target population as whole. So, when samples are drawn only from those who live in households, excluding those who are homeless and live in group quarters, the average income is likely to be higher in those who respond than among those who had no chance to respond. Those in households also are more likely to be married and to have jobs than those without homes. Estimates from those responding to surveys based on households will systematically overestimate the percentage of the whole population that have those characteristics.

One final point to note is that the random variability of sample estimates, the sampling error, and the bias associated with a sample, are not necessarily related at all. If a survey plan systematically leaves out, or underrepresents, some people who are distinctive in ways relevant to the survey's goals, it is quite possible to have a very consistent and stable estimate, with very little sampling error, that is consistently biased and under- or overestimates some characteristics of the population.

ERROR ASSOCIATED WITH ANSWERS

In order to understand error associated with answers to survey questions, one first needs to understand what is being measured. In theory, one could divide

what surveys try to measure into two categories: objective facts and subjective states. Objective facts include a person's height, whether or not a person is employed at a job, and whether or not a person voted in the last election. Subjective states include how much of the time the person has felt tired and whether or not a person has liberal or conservative political views.

Conceptually, the way we assess the answers to a question is to measure how well they correspond to the truth. If we are asking survey respondents about objective facts, such as their height or whether or not they are employed at a job, in theory we could obtain independent information against which to evaluate the answers to the survey question. We could measure the respondent's height; we could verify employment status by looking at records. We could directly assess how accurate the answers were. In contrast, there is no objective way to verify or evaluate a person's report about how much of the time he or she has felt tired.

Psychometricians, those who specialize in the measurement of psychological states, think of answers as consisting of two components: the true score, what a perfect reporter with perfect knowledge would give as an answer, plus some element of error.

$$x_i = t_i + e_i$$

Where

x_i is the answer given by individual i

t_i is the true value for individual i

e_i is the error in the answer given by individual i

Errors can be caused by all kinds of things: misunderstanding the question, not having the information needed to answer, and distorting answers in order to look good are only a few examples. Some respondents might not know their height exactly; others might round their height up or down, thinking that being taller or shorter might be more attractive. Respondents' estimates of how tired they have been over the past week may be affected by how tired they feel at the time they are answering the questions. The point is that to the extent that answers are affected by factors other than the facts on which the answer should be based, there is error in the answer.

Validity is the term that psychologists use to describe the relationship between an answer and some measure of the true score. Looking at the equation above, the goal of the psychometrician and the survey methodologist is to make the error term (e) as small as possible so the answers mainly reflect the true score (or answer).

In a way, the error term in psychometrics is similar to sampling error discussed previously. If the error associated with answers is random, resulting in

answers that err sometimes in one direction, sometimes in another direction, the result is less certainty or confidence in how well the answers are measuring what we want to measure. The larger the value of e (in the equation above), the greater the chance that any individual's answer will be in error. However, across many answers from among individuals, the average answer should be the same as the average true value.

For questions designed to measure objective facts, but not subjective states, there also is the potential that answers will be biased. In a way completely analogous to bias with respect to the samples, what bias means in this context is that on average the errors in answers, the way in which answers differ from the true score, are more likely to be in one direction than another. As examples, respondents on average underreport how many cigarettes they smoke and how much alcohol they drink, while they tend to overstate whether or not they voted. Estimates of these behaviors are likely to be biased—systematically different from the true scores.

The idea of validity for subjective measures cannot be observed directly, but is inferred from studies of how answers are related to other similar measures. The calculations are more complicated, but the end results of estimates of validity are the same: an estimate of how well answers reflect the construct they are designed to measure. (Cronbach & Meehl, 1955; Saris & Andrews, 1991). In contrast, since we cannot directly measure the true value of subjective states, we also cannot measure bias—the degree to which answers systematically differ from a true score in one direction.

RECAPPING THE NATURE OF ERROR IN SURVEYS

Thus, for both of the key inferences that are critical to the quality of survey estimates, the inference that answers can be used to accurately describe a sample of respondents and that we can accurately generalize from a sample of respondents to an entire population, there are two analogous kinds of error: random variability around the true values and systematic (biased) differences between the samples who answer questions and the whole population or between the answers that are given and the true values for those who are answering. Throughout the book, as error is being discussed, readers need to be sure they know which kind of error is meant. Is it the problem of generalization from the sample of respondents to the population or generalization from the answers to the reality we want to describe? Is the concern with respect to error focused on minimizing random variation, variation that may by chance make our samples or our answers different from the true values of

TABLE 2.1
Examples of Error by Types of Error and Type of Inference

	Types of Error	
Inference	*Random*	*Biased*
From sample to population	Sampling error	Example: Those who are over 65 are consistently less likely to respond to telephone surveys, and hence are underrepresented in telephone survey data
From answers to true characteristics	Invalidity	Example: Number of cigarettes smoked is consistently underreported in surveys

the population, or is there some kind of systematic (biasing) error in our data, stemming either from having some elements of the population underrepresented in our sample of respondents or from some systematic distortion of answers to the questions that we posed? Table 2.1 graphically displays four kinds of error that affect our survey estimates.

When trying to evaluate the confidence one can have in estimates based on surveys, it is important to keep in mind all four of these kinds of error. They are different, they usually result from different aspects of the way a survey is executed, and they have different impacts on the ability of a survey to address the questions it is designed to answer.

Given this orientation to the various meanings of *error,* we now proceed to discuss the significance of the way a survey is designed and executed for the confidence one can have in the survey results.

Further Readings

Groves, R. M. (1989). *Survey errors and survey costs*. New York: John Wiley.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology* (Chap. 2). New York: John Wiley.

Lessler, J. T., & Kalsbeek, W. D. (1992). *Nonsampling error in surveys.* New York: John Wiley.