AP Statistics Tutorial: Bias in Survey Sampling

In survey sampling, **bias** refers to the tendency of a sample statistic to systematically over- or under-estimate a population parameter.

Bias Due to Unrepresentative Samples

A good sample is **representative**. This means that each sample point represents the attributes of a known number of population elements.

Bias often occurs when the survey sample does not accurately represent the population. The bias that results from an unrepresentative sample is called **selection bias**. Some common examples of selection bias are described below.

- **Undercoverage**. Undercoverage occurs when some members of the population are inadequately represented in the sample. A classic example of undercoverage is the *Literary Digest* voter survey, which predicted that Alfred Landon would beat Franklin Roosevelt in the 1936 presidential election. The survey sample suffered from undercoverage of low-income voters, who tended to be Democrats.

  How did this happen? The survey relied on a convenience sample, drawn from telephone directories and car registration lists. In 1936, people who owned cars and telephones tended to be more affluent. Undercoverage is often a problem with convenience samples.
- **Nonresponse bias**. Sometimes, individuals chosen for the sample are unwilling or unable to participate in the survey. Nonresponse bias is the bias that results when respondents differ in meaningful ways from nonrespondents. The *Literary Digest* survey illustrates this problem. Respondents tended to be Landon supporters; and nonrespondents, Roosevelt supporters. Since only 25% of the sampled voters actually completed the mail-in survey, survey results overestimated voter support for Alfred Landon.

  The *Literary Digest* experience illustrates a common problem with mail surveys. Response rate is often low, making mail surveys vulnerable to nonresponse bias.
- **Voluntary response bias**. Voluntary response bias occurs when sample members are self-selected volunteers, as in voluntary samples. An example would be call-in radio shows that solicit audience participation in surveys on controversial topics (abortion, affirmative action, gun control, etc.). The resulting sample tends to overrepresent individuals who have strong opinions.

Random sampling is a procedure for sampling from a population in which (a) the selection of a sample unit is based on chance and (b) every element of the population has a known, non-zero probability of being selected. Random sampling helps produce representative samples by eliminating voluntary response bias and guarding against undercoverage bias. All probability sampling methods rely on random sampling.

Bias Due to Measurement Error

A poor measurement process can also lead to bias. In survey research, the measurement process includes the environment in which the survey is conducted, the way that questions are asked, and the state of the survey respondent.

Response bias refers to the bias that results from problems in the measurement process. Some examples of response bias are given below.

- **Leading questions**. The wording of the question may be loaded in some way to unduly favor one response over another. For example, a satisfaction survey may ask the respondent to indicate where she is satisfied, dissatisfied, or very dissatified. By giving the respondent one response option to express satisfaction and two response options to express dissatisfaction, this survey question is biased toward getting a dissatisfied response.
- **Social desirability**. Most people like to present themselves in a favorable light, so they will be reluctant to admit to unsavory attitudes or illegal activities in a survey, particularly if survey results are not confidential. Instead, their responses may be biased toward what they believe is socially desirable.

Sampling Error and Survey Bias

A survey produces a sample statistic, which is used to estimate a population parameter. If you repeated a survey many times, using different samples each time, you would get a different sample statistic with each replication. And each of the different sample statistics would be an estimate for the *same* population parameter.

If the statistic is unbiased, the average of all the statistics from all possible samples will equal the true population parameter; even though any individual statistic may differ from the population parameter. The variability among statistics from different samples is called **sampling error**.

Increasing the sample size tends to reduce the sampling error; that is, it makes the sample statistic less variable. However, increasing sample size does not affect survey bias. A large sample size cannot correct for the methodological problems (undercoverage, nonresponse bias, etc.) that produce survey bias. The *Literary Digest* example discussed above illustrates this point. The sample size was very large - over 2 million surveys were completed; but the large sample size could not overcome problems with the sample - undercoverage and nonresponse bias.

Source: http://stattrek.com/ap-statistics-2/survey-sampling-bias.aspx

Major Sources of Bias in Research Studies

There are two types of error associated with most forms of research: random and systematic. Random errors, i.e., those due to sampling variability or measurement precision, occur in essentially all quantitative studies and can be minimized but not avoided. Systematic errors, or *biases*, are reproducible inaccuracies that produce a consistently false pattern of differences between observed and true values. Both random and systematic errors can threaten the validity of any research study. However, random errors can be easily determined and addressed using statistical analysis; most systematic errors or biases cannot. This is because biases can arise from innumerable sources, including complex human factors. For this reason, avoidance of systematic errors or biases is the task of proper research design

Major Categories of Research Bias

There are many different types of biases described in the research literature. The most common categories of bias that can affect the validity of research include the following:

1. *Selection biases*, which may result in the subjects in the sample being unrepresentative of the population of interest
2. *Measurement biases*, which include issues related to how the outcome of interest was measured
3. *Intervention (exposure) biases*, which involve differences in how the treatment or intervention was carried out, or how subjects were exposed to the factor of interest

More detail on each of these three categories of bias is provided below, including some specific types and examples. For more detail on the various types of biases, including how they can be controlled, link to the following article:

Hartman, J.M., Forsen, J.W., Wallace, M.S., Neely, J.G. (2002). Tutorials in clinical research: Part IV: Recognizing and controlling bias. *Laryngoscope, 112*, 23-31.

Selection Biases

Selection biases occur when the groups to be compared are different. These differences may influence the outcome. Common types of sample (subject selection) biases include volunteer or referral bias, and nonrespondent bias. By definition, nonequivalent group designs also introduce selection bias.

Volunteer or referral bias. Volunteer or referral bias occurs because people who volunteer to participate in a study (or who are referred to it) are often different than non-volunteers/non-referrals. This bias usually, but not always, favors the treatment group, as volunteers tend to be more motivated and concerned about their health.

Nonrespondentbias. Nonrespondent bias occurs when those who do not respond to a survey differ in important ways from those who respond or participate. This bias can work in either direction.

Measurement Biases

Measurement biases involve systematic error that can occur in collecting relevant data. Common measurement biases include instrument bias, insensitive measure bias, expectation bias , recall or memory bias, attention bias, and verification or work-up bias.

Instrument bias. Instrument bias occurs when calibration errors lead to inaccurate measurements being recorded, e.g., an unbalanced weight scale.

Insensitive measure bias. Insensitive measure bias occurs when the measurement tool(s) used are not sensitive enough to detect what might be important differences in the variable of interest.

Expectation bias. Expectation bias occurs in the absence of masking or blinding, when observers may err in measuring data toward the expected outcome. This bias usually favors the treatment group

Recall or memory bias. Recall or memory bias can be a problem if outcomes being measured require that subjects recall past events. Often a person recalls positive events more than negative ones. Alternatively, certain subjects may be questioned more vigorously than others, thereby improving their recollections.

Attention bias. Attention bias occurs because people who are part of a study are usually aware of their involvement, and as a result of the attention received may give more favorable responses or perform better than people who are unaware of the study's intent.

Verification or work-up bias. Verification or work-up bias is associated mainly with test validation studies. In these cases, if the sample used to assess a measurement tool (e.g., diagnostic test) is restricted only to who have the condition of factor being measured, the sensitivity of the measure can be overestimated.

Intervention (Exposure) Biases

Intervention or exposure biases generally are associated with research that compares groups. Common intervention biases include: contamination bias, co-intervention bias, timing bias(es), compliance bias, withdrawal bias, and proficiency bias.

Contamination bias. Contamination bias occurs when members of the 'control' group inadvertently receive the treatment or are exposed to the intervention, thus potentially minimizing the difference in outcomes between the two groups.

Co-intervention bias. Co-intervention bias occurs when some subjects are receiving other (unaccounted for) interventions at the same time as the study treatment.

Timing bias(es). Different issues related to the timing of intervention can bias. If an intervention is provided over a long period of time, maturation alone could be the cause for improvement. If treatment is very short in duration, there may not have been sufficient time for a noticeable effect in the outcomes of interest.

Compliance bias. Compliance bias occurs when differences in subject adherence to the planned treatment regimen or intervention affect the study outcomes..

Withdrawal bias. Withdrawal bias occurs when subjects who leave the study (drop-outs) differ significantly from those that remain.

Proficiency bias. Proficiency bias occurs when the interventions or treatments are not applied equally to subjects. This may be due to skill or training differences among personnel and/or differences in resources or procedures used at different sites.

Source: http://www.umdnj.edu/idsweb/shared/biases.htm

Sampling

Main Contents page | Index of all entries

STEPS Target Population

The target population is the entire group a researcher is interested in; the group about which the researcher wishes to draw conclusions.

Example

Suppose we take a group of men aged 35-40 who have suffered an initial heart attack. The purpose of this study could be to compare the effectiveness of two drug regimes for delaying or preventing further attacks. The target population here would be all men meeting the same general conditions as those actually included in the study.

STEPS Matched Samples

Matched samples can arise in the following situations:

a. Two samples in which the members are clearly paired, or are matched explicitly by the researcher. For example, IQ measurements on pairs of identical twins.
b. Those samples in which the same attribute, or variable, is measured twice on each subject, under different circumstances. Commonly called repeated measures. Examples include the times of a group of athletes for 1500m before and after a week of special training; or the milk yields of cows before and after being fed a particular diet.

Sometimes, the difference in the value of the measurement of interest for each matched pair is calculated, for example, the difference between before and after measurements, and these figures then form a single sample for an appropriate statistical analysis.

STEPS Independent Sampling

Independent samples are those samples selected from the same population, or different populations, which have no effect on one another. That is, no correlation exists between the samples.

STEPS Random Sampling

Random sampling is a sampling technique where we select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has a known, but possibly non-equal, chance of being included in the sample.

By using random sampling, the likelihood of bias is reduced.

Compare simple random sampling.

STEPS Simple Random Sampling

Simple random sampling is the basic sampling technique where we select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample. Every possible sample of a given size has the same chance of selection; i.e. each member of the population is equally likely to be chosen at any stage in the sampling process.
Compare random sampling.

STEPS Stratified Sampling

There may often be factors which divide up the population into sub-populations (groups / strata) and we may expect the measurement of interest to vary among the different sub-populations. This has to be accounted for when we select a sample from the population in order that we obtain a sample that is representative of the population. This is achieved by stratified sampling.
A stratified sample is obtained by taking samples from each stratum or sub-group of a population.
When we sample a population with several strata, we generally require that the proportion of each stratum in the sample should be the same as in the population.
Stratified sampling techniques are generally used when the population is heterogeneous, or dissimilar, where certain homogeneous, or similar, sub-populations can be isolated (strata). Simple random sampling is most appropriate when the entire population from which the sample is taken is homogeneous.
Some reasons for using stratified sampling over simple random sampling are:

a. the cost per observation in the survey may be reduced;
b. estimates of the population parameters may be wanted for each sub-population;
c. increased accuracy at given cost.

Example
Suppose a farmer wishes to work out the average milk yield of each cow type in his herd which consists of Ayrshire, Friesian, Galloway and Jersey cows. He could divide up his herd into the four sub-groups and take samples from these.

STEPS Cluster Sampling

Cluster sampling is a sampling technique where the entire population is divided into groups, or clusters, and a random sample of these clusters are selected. All observations in the selected clusters are included in the sample.
Cluster sampling is typically used when the researcher cannot get a complete list of the members of a population they wish to study but can get a complete list of groups or 'clusters' of the population. It is also used when a random sample would produce a list of subjects so widely scattered that surveying them would prove to be far too expensive, for example, people who live in different postal districts in the UK.
This sampling technique may well be more practical and/or economical than simple random sampling or stratified sampling.
Example
Suppose that the Department of Agriculture wishes to investigate the use of pesticides by farmers in England. A cluster sample could be taken by identifying the different counties in England as clusters. A sample of these counties (clusters) would then be chosen at random, so all farmers in those counties selected would be included in the sample. It can be seen here then that it is easier to visit several farmers in the same county than it is to travel to each farm in a random sample to observe the use of pesticides.

STEPS **Quota Sampling**

Quota sampling is a method of sampling widely used in opinion polling and market research. Interviewers are each given a quota of subjects of specified type to attempt to recruit for example, an interviewer might be told to go out and select 20 adult men and 20 adult women, 10 teenage girls and 10 teenage boys so that they could interview them about their television viewing.
It suffers from a number of methodological flaws, the most basic of which is that the sample is not a random sample and therefore the sampling distributions of any statistics are unknown.

STEPS **Spatial Sampling**

This is an area of survey sampling concerned with sampling in two (or more) dimensions. For example, sampling of fields or other planar areas.

STEPS **Sampling Variability**

Sampling variability refers to the different values which a given function of the data takes when it is computed for two or more samples drawn from the same population.

STEPS **Standard Error**

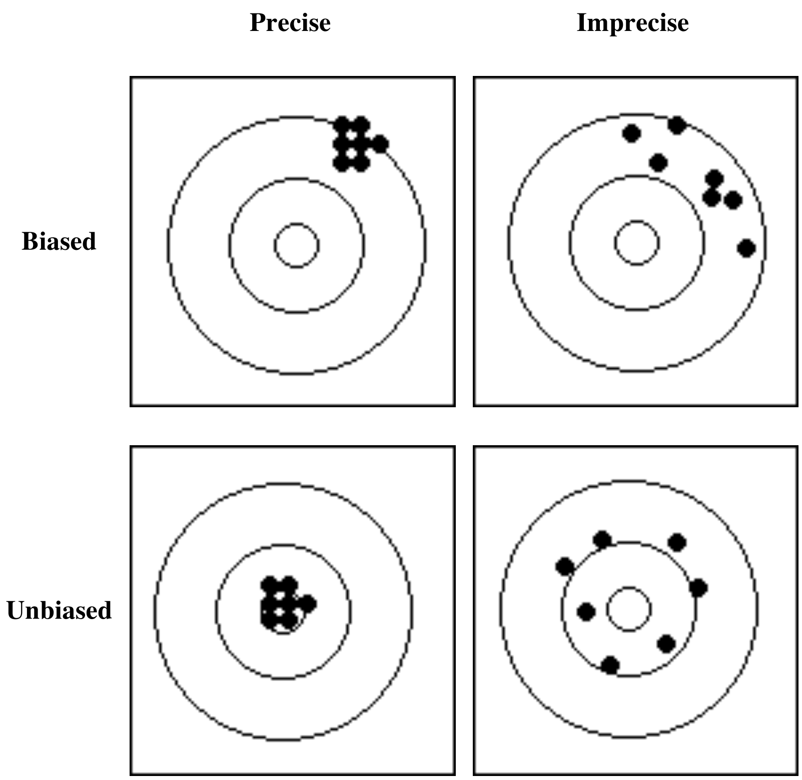
Standard error is the standard deviation of the values of a given function of the data (parameter), over all possible samples of the same size.

Source: http://www.stats.gla.ac.uk/steps/glossary/sampling.html

Bias

Bias is a term which refers to how far the average statistic lies from the parameter it is estimating, that is, the error which arises when estimating a quantity. Errors from chance will cancel each other out in the long run, those from bias will not.

The following illustrates bias and precision, where the target value is the bullseye:



Example

The police decide to estimate the average speed of drivers using the fast lane of the motorway and consider how it can be done. One method suggested is to tail cars using police patrol cars and record their speeds as being the same as that of the police car. This is likely to produce a biased result as any driver exceeding the speed limit will slow down on seeing a police car behind them. The police then decide to use an unmarked car for their investigation using a speed gun operated by a constable. This is an unbiased method of measuring speed, but is imprecise compared to using a calibrated speedometer to take the measurement.

See also precision.

Precision

Precision is a measure of how close an estimator is expected to be to the true value of a parameter. Precision is usually expressed in terms of imprecision and related to the standard error of the estimator. Less precision is reflected by a larger standard error.

See the illustration and example under bias for an explanation of what is meant by bias and precision.