Sullivan Section 1.4

Gene Quinn

Sources of Error in Sampling

Sampling is usually necessary because even if the entire population in an experiment or study can be identified, it is too costly to survey everyone in that population.

The sampling process itself introduces a number of problems, some of which can be remedied, and some which have no solution.

Sources of Error in Sampling

Sampling is usually necessary because even if the entire population in an experiment or study can be identified, it is too costly to survey everyone in that population.

The sampling process itself introduces a number of problems, some of which can be remedied, and some which have no solution.

Sources of error fall into two major categories:

- **Nonsampling errors** are errors that result from the survey process and generally cannot be fixed.
- **Sampling errors** are errors that result from using a sample to represent the entire population. Sampling errors can be controlled by careful design of the experiment.

Nonsampling Errors

Definition: Nonsampling errors are errors that result from the survey process.

nonsampling errors include:

- using an incomplete or inappropriate frame
- nonresponse by some individuals selected in the sample
- poorly trained interviewers
- data entry and recording errors
- poor questionnaire design
- poorly worded questions

Nonsampling Errors

Definition: Nonsampling errors are errors that result from the survey process.

nonsampling errors include:

- using an incomplete or inappropriate frame
- nonresponse by some individuals selected in the sample
- poorly trained interviewers
- data entry and recording errors
- poor questionnaire design
- poorly worded questions

Nonsampling errors generally have no solution.

Sampling Errors

Definition: Sampling errors are errors that result from using a sample to represent the entire population.

Sampling Errors

Definition: Sampling errors are errors that result from using a sample to represent the entire population.

Sampling error can be controlled by:

- carefully designing the experiment
- choosing an appropriate sampling method
- using an appropriate sample size

Sources of Error - the Frame

One source of error is the frame, or list of all individuals in the population.

Sources of Error - the Frame

One source of error is the frame, or list of all individuals in the population.

In some cases, the frame may be relatively easy to obtain.

For example, in many samples related to election polls, the list of registered voters often provides a good frame.

Sources of Error - the Frame

One source of error is the frame, or list of all individuals in the population.

In some cases, the frame may be relatively easy to obtain.

For example, in many samples related to election polls, the list of registered voters often provides a good frame.

For studies of the population as a whole, it may be very difficult to obtain

a frame that does not omit certain subgroups.

Sources of Error - Nonresponse

A source of error that is nearly always present is **nonresponse**.

This occurs when an individual selected for the sample fails to respond for some reason, either intentionally or because they cannot be reached.

Sources of Error - Nonresponse

A source of error that is nearly always present is **nonresponse**.

This occurs when an individual selected for the sample fails to respond for some reason, either intentionally or because they cannot be reached.

Techniques to reduce the number of nonrespondents include one or more attempts to contact them. These may be in the form of follow-up phone calls, or visits to the homes of nonrespondents to contact them in person.

Sources of Error - Nonresponse

A source of error that is nearly always present is **nonresponse**.

This occurs when an individual selected for the sample fails to respond for some reason, either intentionally or because they cannot be reached.

Techniques to reduce the number of nonrespondents include one or more attempts to contact them. These may be in the form of follow-up phone calls, or visits to the homes of nonrespondents to contact them in person.

Another technique is to offer incentives for respondents in the form of

cash payments, or entry into a lottery, or other rewards for responding.

Sources of Error - Interviewers

The skills of the people who conduct the interviews are often important.

Many surveys include questions of a personal or embarassing nature, and it is important for the interviewer to communicate with the respondent in a way that will make them comfortable giving truthful answers.

Sources of Error - Data Checks

Another type of error occurs when data is recorded or entered incorrectly.

The recording and data entry operations should be validated to ensure

correctness.

Sources of Error - Questionnaire Design

A questionnaire must be carefully designed to produce accurate results.

Questions may be **open**, which allow the respondent to answer freely, or **closed**, which present the respondent with a list of responses from which one must be chosen.

Sources of Error - Questionnaire Design

A questionnaire must be carefully designed to produce accurate results.

Questions may be **open**, which allow the respondent to answer freely, or **closed**, which present the respondent with a list of responses from which one must be chosen.

The wording of questions is also important. Questions must not be worded in a way that encourages or discourages certain answers.

Sources of Error - Questionnaire Design

A questionnaire must be carefully designed to produce accurate results.

Questions may be **open**, which allow the respondent to answer freely, or **closed**, which present the respondent with a list of responses from which one must be chosen.

The wording of questions is also important. Questions must not be worded in a way that encourages or discourages certain answers.

The order in which questions are asked can also impact the results of a survey, as well as the order in which alternative responses are presented.