Ch 4 Sampling errors

sampling errors: caused by the act of taking a sample. They cause sample results to be different from the results of a census

Random Sampling error: the deviation between the sample statistic and the population parameter caused by change in selecting a random sample.

The margin of error in a confidence statement includes only random sampling error.

Non-sampling Error: errors not related to the act of selecting a sample from the population. They can be present even in a census.

Undercoverage: occurs when some groups in the population are left out of the process of choosing a sample.

Questions to Ask Before Believing a Poll.

Who carried out the survey? Any personal gain from outcome?

What was the population? Whose opinions were being asked?

How was the sample selected? Look for mention of random sampling.

How large was the sample? Even better, find out both the sample size and margin of error within which results of 95% of all samples drawn as this one was would fall.

What was the response rate? what percent of the original subjects actually provided information

How were the subjects contacted? Phone, mail, face to face

When was the survey conducted? Was it just after some event that might have influenced opinion?

What were the exact questions asked?

Nonsampling Errors

Nonsampling errors can be present in a census

Processing error: mistakes in mechanical tasks such as imput.

Response error: when a subject gives an incorrect response (lies/doesn't understand)

Nonresponse: failure to obtain data from an individual selected for a sample. Most nonresponse happens because subjects can't be contacted or because some subjects contacted refuse to cooperate

What the Margin of error doesn't say:

covers only random sampling error.

nonresponse, under coverage and other practical difficulties can cause large bias that is not covered by margin of error.

How to Live with Non sampling Errors

First: substitute other households for non responders.

Replacing nonresponders with other households in the same neighbourhood may reduce bias.

Once all the data is in, weight the responses.

Weighting is an attempt to correct sources of bias. For example, if there are too many households that did not respond, the survey gives weight to the houses that did respond.

Wording Questions

question wording may influence the results of a sample survey

If the questions are slanted to favor one response over others, we have another source of non sampling error.

Taking the sample in several stages with clusters at the final stage saves time for interviewers grouping the sample households first in PSUs and then in clusters.

Wording Questions (cont)

Stratified Sample:

step 1: divide the sampling frame into distinct groups of individuals, called strata.

step 1.2: choose the strata according to any special interest you have in certain groups within the population or because the individuals in each stratum resemble each other.

step 2: take a separate SRS in each stratum and combine these to make up the complete sample.

Probability Sample: chosen by chance. Know what samples are possible and what chance/probability each possible sample has.

Some probability samples, such as stratified samples, don't allow all possible samples from the population and may not give an equal chance to all the samples they do allow.