NONSAMPLING ERRORS

Eugene Sauls, University of Arkansas
and
James A. Millar, University of Arkansas

Introduction

There are two classes of nonsampling errors which occur in a survey. These classes are herein designated as nonresponsive errors and improper response errors. Nonresponsive errors arise when an inquiry is not answered. Improper response errors arise when a respondent offers an incorrect answer to an inquiry.

This paper reports on a study which investigated the existence of each of these classes of errors. The study was conducted with the use of deposit accounts in a credit union and loan accounts in a bank. Confirmations of the correctness of account balances were requested from the depositors and debtors after the account balances were adjusted.

The study shows that nonresponse and improper errors do exist and must be recognized by surveyors.

Nonresponses

Nonresponse errors occur because the subject is unable to respond or because the subject refuses to answer. The subject may be unable to respond because of a lack of knowledge or because of an inability to understand the inquiry. The refusal may result because the subject views the inquiry as an unjustifiable invasion of privacy or because the subject simply will not take the time to respond.

The subject may be asked questions for which he lacks the knowledge to answer. In some instances the lack of knowledge may be a function of background characteristics. For example the nonresponsive rate to a question concerning educational philosophies may be higher among less educated subjects. If the lack of knowledge is a function of the characteristics of the subject, a bias is injected into the results. This bias may be compensated for by stratifying the population by its relevant characteristics. In some cases the subjects favoring a particular position may become more knowledgeable because they seek information supporting their position. This form of bias cannot be compensated for; therefore, conclusions apply only to the population to which the subjects belong. For example, one might expect a higher nonresponse rate from athiest on a question concerning support for the divinity of Jesus Christ than would be expected from Christians. The degree of devoutness would presumably affect the nonresponse rate. If the population cannot be stratified according to its characteristics or the ratio of each strata to the population is unknown, an immeasurable bias is present.

The interviewer may pose the question in such a way that the subject is not certain of the questions. Rather than attempt to interpret the question, the subject may choose to not respond. If the question is worded so that any reservation as to the question would not be a function of the characteristics of the subject, no bias is present.

The subject may refuse to respond on the grounds that the question is an unjustifiable invasion of privacy. There would generally be no reason to believe that those who refused to respond for this reason would be any different from those who responded.

A subject's view of his time constraints may result in nonresponse errors. A subject who felt that he was busy or that his time is too valuable to be spent answering questionnaires may refuse to respond. Perhaps a bias is frequently encountered because of the subject's perceived time constraint. For example, public officials and business executives would probably be much more inclined to not respond because of the time involved than would a trademan or a laborer.

Improper Responses

Improper response errors occur because the respondent (1) does not know the correct answer but offers one nevertheless, (2) does not understand the question and therefore misinterprets the question, (3) is influenced by the interviewer the offer the answer perceived by the subject as the answer most suitable to the interviewer or (4) deliberately misstates his position in order to influence others.

Opinion surveys have a relatively high probability of being made of subjects who do not know the answer. For example, a subject may respond that blue is his favorite color and he may actually believe he is telling the truth whereas a thorough analysis of his preference through choice situations may reveal that red is his favorite color. In most cases, however, the lack of knowledge is probably not a function of some characteristic of the population; therefore, there would generally be no significant bias in the overall results of a large sample. Technical questions may result in a serious bias because some groups of the population do not have sufficient knowledge to answer but answer anyway.

Misinterpretation of a question poses a serious problem. There might be many instances where the interpretation of the question is a function of some characteristic of the population, e.g., education.

Prior Studies

Robert Ferber reported the results of a study in which interviews were conducted with 188 urban subjects and 33 farm subjects.[1] The subjects were queried about their financial holdings. The balances of certain deposit accounts held by the subjects were known. The nonresponse rate was seventy-nine percent of subjects contacted. More importantly, non-responders had a significantly higher average balance than responders. Population estimates would be significantly below the true balance in cases where this condition existed.

Research Design

In the financial audits of organizations it is standard practice for the auditors to circularize the accounts, i.e., confirmation requests are sent to creditors and debtors asking them to confirm the balance or to report a difference. The evidence gathered through this procedure assists the auditor in arriving at a conclusion as to the reasonableness of the account balance.

Confirmation requests were sent under routine audit conditions to samples of deposit accounts of the MSU Employees Credit Union and loan accounts of the Continental Illinois National Bank and Trust Company of Chicago.

Deposit Accounts

The experiment was conducted on the 456 time deposit accounts as of February 29, 1968. Twenty-two accounts were deleted for various reasons. Certificates are issued to the depositors, thus there was tangible evidence on the account balance. The balance of the account represented principal only and averaged approximately $3,000.

Three samples were drawn: K_1 - a sample of fifty accounts which were circularized without adjusting the balances of the accounts. This sample was the control group.

K_2 - a sample of thirty accounts, the balances of which were adjusted by a positive adjustment of ten percent - (rounded to the nearest fifty dollars) with a maximum adjustment of $500. One recipient in this sample learned that the confirmation request was part of a study; therefore, that account was deleted from the sample, leaving twenty-nine accounts in the sample.

K_3 - a sample of thirty accounts, the balances of which were adjusted by negative ten percent (rounded to the nearest fifty dollars) with maximum adjustment of $500.

Second requests were mailed fourteen days after the first request to those who did not respond to the first requests.

Loan Accounts

Direct personal loan accounts were used in this part of the study. The average balance of the personal loan accounts in the sample was approximately $1,200 and the average balance of the automobile loan accounts was approximately $1,800. The experiment was conducted during March and April 1968. Two samples were drawn:

b_1- a sample of one hundred accounts which were circularized without adjusting the balances of the accounts. This sample represented the control group.

b_2- a sample of thirty accounts, the balances of which were adjusted by a positive adjustment of approximately ten percent.

Second requests were mailed to sample b_1 about two weeks after the first requests were mailed. Second requests were not sent to sample b_2 because of the adverse reaction to the first requests.

Hypotheses

Disregarding the effects on statistical models, the effects of nonresponse and improper response errors on conclusions drawn from questionnaires are a function of the extent of errors and the degree of errors. The extent and degree of errors varies with the circumstances. There can be no universal criteria for acceptable error rates. In order to provide test criteria, arbitrary error rates were set.

The test criterion for proper responses was 0.70 when there was no effort to follow up on nonresponses. When second requests were sent to nonresponsers to the first requests, the test criterion was 0.90 (approximately = 0.70 + (0.70) x (0.30). Improper response errors were tested against a 0.05 test criterion. Nonresponse errors were tested against the nonresponse error rate in the control group.

The hypotheses are designated as "D" for deposits for "L" for loans. The second identifier is: "P" for improper response or "N" for nonresponse. The third identifier is the sequential number of the hypothesis.

H_0:DP1 - The proportion proper responses to overstated amounts is 0.90.

H_a: The proportion of proper responses to overstated amounts is less than 0.90.

H_0:DP2 - The proportion of proper responses to understated amounts is 0.90.

H_a: The proportion of proper responses to understated amounts is less than 0.90.

H_0D1 - The proportion of improper responses to overstated amounts is 0.05.

H_a: The proportion of improper responses to overstated amounts is more than 0.05.

H_0:D14 - The proportion of improper responses to understated amounts is 0.05.

H_a: The proportion of improper responses to overstated amounts is more than 0.50.

H_0D15 - The proportion of improper responses to overstated amounts is equal to that of understated amounts.

H_a: The proportion of improper responses to overstated amounts is not equal to that of understated amounts.

H_0DN6 - The proportion of nonresponses to incorrect amounts is equal to that of correct amounts.

H_a: The proportion of nonresponses to incorrect amounts is not equal to that of correct amounts.

H_0DN7 - The proportion of nonresponses to overstated amounts is equal to that of understated amounts.

H_a: The proportion of nonresponse to overstated amounts is not equal to that of understated amounts.

H_0DP2 - The proportion of proper responses to understated amounts is 0.90.

H_a: The proportion of proper responses to understated amounts is less than 0.90.

H_0:D13 - The proportion of improper responses to overstated amounts is 0.05.

H_a: The proportion of improper responses to overstated amounts is more than 0.05.

H_0:D14 - The proportion of improper responses to understated amounts is 0.05.

H_a: The proportion of improper responses to overstated amounts is more than 0.50.

H_0:D15 - The proportion of improper responses to overstated amounts is equal to that of understated amounts.

H_a: The proportion of improper responses to overstated amounts is not equal to that of understated amounts.

H_0:DN6 - The proportion of nonresponses to incorrect amounts is equal to that of correct amounts.

H_a: The proportion of nonresponses to incorrect amounts is not equal to that of correct amounts.

H_0DN7 - The proportion of nonresponses to overstated amounts is equal to that of understated amounts.

H_a: The proportion of nonresponse to overstated amounts is not equal to that of understated amounts.

H_0DP8 - The proportion of proper responses to first request on incorrect amounts is equal to that of second requests.

H_a: The proportion of proper response to first requests on incorrect amounts is not equal to that of second request.

H_0DP9 - The proportion of proper responses is independent of the age of the recipient.

H_a: The proportion of proper responses is not independent of the age of the recipient.

H_0DP10 - The proportion of proper responses is independent of the size of amount.

H_a: The proportion of proper responses is not independent of the size of the amount.

H_0LP11 - The proportion of proper responses to overstated amount is 0.70.

H_a: The proportion of proper responses to overstated amounts is less than 0.70.

H_0LI12 - The proportion of improper responses to overstated amounts is 0.05.

H_a: The proportion of improper responses to overstated amounts is greater than 0.05.

H_0LN13 - The proportion of nonresponses to overstated amounts is equal to that of correct amounts.

H_a: The proportion of nonresponses to overstated amounts is not equal to that of correct amounts.

H_0DLP14 - The proportion of proper responses to incorrect deposit accounts is equal to that of loan accounts.

[1]Robert Ferber, "The Reliability of Consumer Surveys on Financial Holdings: Time Deposits," Journal of the American Statistical Association (March 1965), 148-163.

Test Results

Hypothesis	Test Statistic	Decision
1	*Z=.00106	reject
2	*Z=.17361	Not reject
3	*Z=.00184	reject
4	*Z=.18630	Not reject
5	**t= 1.06	Not reject
6	**t= .74	Not reject
7	**t= .05	Not reject
8	**t= .10	Not reject
9	***F= .200	Not reject
10	***F= 1.06	Not reject
11	****Z= 3.19	reject
12	****Z= 1.25	Not reject
13	X^2 = .17	Not reject
14	***F= 7.16	reject
15	***F= 1.57	Not reject
16	***F= 4.64	reject

*Hypergeometric

**Bennett and Franklin Test

	.05	.01
t	2.77	3.64

***Bartletts Transformation

	.05
F	4.46

****Binomial