B4-8-M

SURVEY PRINCIPLES AND TECHNIQUES*

W. T. Federer
Cornell University

I. INTRODUCTION

I like to look upon statistical methods, in the modern sense, as a collection of tools and gadgets which are, sometimes, extremely useful to the experimental worker. These tools, just like the tools of a garage mechanic or a carpenter, are designed for specific purposes. The application of the wrong tool for a specific job results in partial or complete failure. For example, if we wish to drive a nail into a board we most assuredly would not use a screwdriver. Modern statistical methodology is the study of the application of statistical tools to specific problems. The wrong application can sometimes give very misleading answers.

Statistical tools have an important place in evaluation studies of extension programs. One example of the use of statistics in such programs has been reported by the Iowa State College extension staff in a study on the use of a caravan in Cass County, Iowa. The statistical gadget that was useful in this case was the probability sample design. Before discussing the Cass County Survey, let's consider the various aspects of a survey.

II. CONDUCT OF A SURVEY

In the conduct of any survey there are five distinct but closely related stages. Chronologically they are:

(i) Initial planning

(ii) Selection of the sample design

- - - - - - - - - - -

*Talk given before Cornell Extension Club, May 1, 1950

(iii) Drafting the questionnaire

(iv) Operational mechanics of the survey

(v) Tabulation, analysis and publication of results

If any one of these stages is weak, the results of the entire survey will be affected. The old adage that "a chain is not stronger than its weakest link" is applicable here.

(i) Initial Planning

The first step in any survey should be a statement of the objectives. Certain specific questions should be asked and investigated for their practicality and usefulness. For his own good, anyone who plans to take a survey should write out the questions to be answered. He then should consider procedures to follow for the various types of answers that are possible. If there is only one possible answer either in fact or in our own minds, then there is no need to take a survey. Charles F. Sarle of the Bureau of Agricultural Economics and an alumnus of Cornell stated that "a true scientist never sets out to 'prove' his ideas; he tests them. He makes predictions and then conducts experiments to see how well they are verified." Therefore, even if we have certain prejudices, we should be unprejudiced in the conduct of a survey.

Another important phase in the initial planning of a survey is to consider the amount of money and personnel that are available for the study. If either or both are too limited, then we can do nothing more than wait until available resources are at our disposal. If relatively little money and time are available, we would be better off to do one job well rather than do several poorly. Also, if a survey is really important, then the administration should be willing to let some of the other duties be delayed or omitted.

(ii) Selection of the Sample Design

If our population of interest were sufficiently small it would be possible to include every member of the population in our survey. Populations usually are not small enough to enumerate completely. Therefore, it would be impractical and sometimes less accurate to take a complete census. (I doubt,though,if a census ever is complete.) We are forced to include only a part or a sample of the population. The next question is how this sample should be drawn. First of all, it should have the characteristics that it is representative and that it gives an unbiased estimate of the true population value. Many conductors of surveys "think" that they are capable of selecting the individuals who are "representative" of the population. Frankly, I don't believe the human being lives who is capable of selecting individuals who are representative of a population. If one of you has such extraordinary powers, both the Republicans and Democrats would be extremely interested, and it won't take you long to become a millionaire. You could save a lot of money for the Census Bureau, the Bureau of Agricultural Economics, National Analysts, Inc., Gallup, Roper, etc.

Therefore, since most of us are ordinary people and do not possess supernatural powers, we must resort to the methods of Probability Sampling in order to have samples which are both representative and unbiased. A probability sample is one selected in such a way that each element in the population has a known probability, greater than zero in the mathematical sense, of being in the sample. This is in contrast to judgment or non-probability samples wherein the sampling units have unknown probabilities of selection. For example, one type of survey practiced over a period of time in various parts of the country was to decide that an area, such as a township, was "representative" of the population of interest and then to start down the road and enumerate 100 farms. If a farmer refused to answer, this fact was ignored, and the next farmer who would answer was enumerated. These areas are not representative of the county for all characteristics enumerated, although they are considered to be "homogeneous" groups. It is assumed that "birds of a feather flock together." This is only partially true for farm characteristics, marital status, religious preferences, etc. If we are interested in homogeneous groups, the groups of interest should be defined and then precautions should be taken to have homogeneous elements within the groups.

If step (i) above were observed rigidly, the experimenter would not run into the difficulties mentioned. In the above example the objectives have not been defined clearly. Surveys have been conducted in this manner for the last twenty to thirty years. The <u>relative</u> utility of this type of data has never been appraised.

In designing a sample five decisions must be made:

(1) The size of the universe or population and the units of observation must be defined. The geographic limits usually must be defined as well as the elements within the defined area. The elements of interest within the area may be only, say, dairy farmers. All other elements are excluded from our defined universe within the defined area.

The unit of observation is one on which separate schedules or data are obtained. It may be a farm, or a group of farms, an individual, or a group of individuals, etc.

(2) It is necessary to set the size of sample. This is most commonly determined by the available funds. The problem then is to select a sample design which will yield the most information per dollar spent. If the cost is not the limiting factor, some other limiting factors might be the degree of precision desired, available personnel, kinds of estimates to be made, etc.

It may be that the experimenter could send questionnaires by mail to everyone in the universe. However, if the return was non-representative, as it is in nearly all mailed questionnaires, the data, alone, would be relatively worthless. The cost of mailing and processing the schedules would be wasted. Since the above usually holds true, the experimenter would be much better off to take a relatively small probability sample. In this way he would obtain more information per dollar spent. It might be well to point out here that the Census Bureau knows that it would be able to obtain more accurate information from a sample and with less cost than is obtained with a census. However, they would have to first convince Congress before the costly Census could be eliminated.

(3) One must decide on the restrictions to be employed in the sample design. An example will best illustrate this. Suppose that we wish estimates for the Northeast. Our first step might be to divide the area into regions and then further subdivide the regions into subareas, say counties. The counties would be partitioned into segments. In a probability sample we make a random selection of counties (primary sampling units) within regions and then randomly select the

specified number of segments (secondary sampling units) within the selected counties. The above method of subdividing the total universe into subareas is called stratification or grouping. The purpose of this grouping is to divide the total universe into relatively homogeneous groups; if the stratification is efficiently done, the precision of the estimate is greatly increased.

(4) One must consider the non-response problem as to the number of callbacks the enumerators are to make and as to the effect of non-response on the final estimate. In mailed questionnaires the response will be between 10-30 percent for most surveys. In a more extreme case let's consider that the questionnaire had only questions of interest to dairymen, but this fact escaped your attention and 99 percent of your responses were from dairymen. You can see what this does to your overall estimate.

(5) The experimenter should consider the methods of expansion to area totals. In some very simply designed samples this decision needs little consideration. Other types of sample designs might involve quite complicated methods of expansion. By way of example, proper methods of expansion probably would have prevented the grossly misleading estimates obtained in the Literary Digest.

Most of you will need aid in drawing your sample. Professor A. A. Johnson was in need of such aid when he obtained a probability sample of seed dealers. Professor P. J. McCarthy and I consulted with him on this. The Cass County Study in Iowa required the aid of the Statistical Laboratory. If you do need such help, I suggest that you contact a statistician. We here at Cornell may be able to help you. Due to the excess of consulting, I won't promise anything, but will say that we'll try. I would suggest that you contact Prof. McCarthy first and if no response, try Douglas S. Robson or me in the Biometrics Unit.

(iii) Drafting the Questionnaire

The design and construction of the questionnaire is important because:

(1) If the questions are not clearly stated and properly set up, one will not get the information desired.

(2) A well-designed questionnaire may eliminate the decision one must make about non-response.

(3) If the questionnaire is properly designed, the accuracy and representativeness of the data will be increased.

(4) A properly designed questionnaire simplifies and speeds up the processing of the schedules.

Little else will be said on this very important item except to point out a few precautions to take. First of all, compare your list of objectives in (i) with the questionnaire. You may find that they are only remotely related. This, then, would mean that a revision of the questionnaire is imperative. The facts of interest must be included on the questionnaire. Secondly, do not include irrelevant questions. People are already questioned too much. Thirdly, use as short a schedule as possible. Do not tire or bore your respondents. Fourthly, always pretest the questionnaire.

It would be desirable to let an expert such as O. F. Larsen in Rural Sociology look over your questionnaire before it is to be pretested. If this is not possible, then have someone in another field scrutinize it. For example, Prof. A. A. Johnson and L. D. Kelsey had me look over their questionnaires and their list of objectives. I was above to make several suggested changes. In fact some of your fellow workers, wives or husbands can often make valuable suggestions on a questionnaire before you pretest it or send it to one of the "experts."

(iv) Operational Mechanics

Sample surveys or censuses that are conducted by the experimenter himself or with one or two assistants do not offer much difficulty in the way of operational mechanics. However, as soon as the survey becomes larger several operational problems begin to "rear their ugly heads."

In a large survey the next step we have to take after steps (i), (ii), and (iii) listed above is to hire enumerators. These enumerators must then be trained and transported to the place of operation. The schedules taken by the enumerators must homehow be returned to the central office. Also, we must devise means of checking on our enumerators. Then there is the payment of salaries, travel, and lodging for these people. As one can see, the larger and more extensive the survey, the more important the operational mechanics of a survey become.

To try to impress upon you the importance of adequate training of enumerators and of adequate checks, let me relate two incidents that happened to a fellow statistician, say Mr. F, when he was enumerated in two different surveys.

In the first instance a young lady came to Mr. F's home and told him that she was taking a survey for a laxative company. After the proper introductions, she asked a number of questions. Among them was one, "How often do you use laxatives?" (This was a loaded question.) Mr. F answered that he never had. Then to his horror he observed that she wrote down that he used laxatives 4-5 times a week. Mr. F asked why she had deliberately written down a falsehood and she replied that the company wouldn't like it if she wrote down his answer. Mrs. F answered in the same manner as her husband, but the enumerator indicated on the schedule that Mrs. F used laxatives every day of the week.

In a second survey in which Mr. F happened to be one of the "unlucky" individuals drawn, he was asked why he didn't shop in Ithaca. Mr. F, being a very serious individual, explained in detail why he didn't shop in Ithaca. Again the enumerator did not write down his answers but something else. When asked by the indignant Mr. F why he did not write down the correct answer, the enumerator replied that it really didn't matter because everything would average out anyway.

Thus, as you can see, proper training and checking of enumerators is imperative. For the larger surveys, I would suggest the use of some of the experimental designs used in biology as an aid to the experimenter in checking and comparing the results from the individual enumerators.

(v) Tabulation, Analysis, and Publication of Results

The successful conclusion of a survey results in the publication of results. Publication could be in a printed, mimeographed, dittoed, or typed form. A letter might be sufficient to convey the results of a study to the interested party or parties.

Prior to the publication we must consider tabulation of the material and then the statistical analysis. The interpretation of the results should be given in the report. Some helpful hints in the analysis might be:

1. Make up the schedule or questionnaire in such a way as to facilitate tabulating and summarizing the material.

2. The questionnaire might be put on McBee Cards.

3. If the material is to be put on punched cards, be certain to code the material on the questionnaire. If this is done, only one schedule requires coding rather than all of them. For more detailed instructions on this it is suggested that you contact P. J. McCarthy. Remember that coding errors in some surveys may be larger than the actual sampling errors!

4. Consult a statistician for short cuts in analyses of data.

5. Do not carry more than the meaningful number of significant figures at any stage in the analysis.

6. A graph may be more instructive and easier to understand than a long and involved table of figures.

7. Consult a journalist or English teacher for aid in writing your report. You may be surprised at the help they are able to offer if interested. It is too bad that every experiment station doesn't have a consulting journalist.

III. ERRORS IN A SURVEY

Most of the errors in a survey have already been discussed but are brought together here. Some of the sources of errors are:

1. Faulty objectives.
2. Faulty questionnaire construction.
3. Enumerator errors and biases.
4. Non-response.
5. Coding and transcribing errors.
6. Other clerical errors.
7. Faulty statistical analyses.
8. Faulty interpretation
9. Random sampling fluctuations.

Of course there may be others but the above list should include most of them. In many surveys, P. C. Mahalanobis, a well-known Indian statistician, has found that the sampling errors in a survey may be less than one-fourth of the total errors. Therefore, he has been concentrating on means of reducing the other sources of errors in surveys.

IV. AN EXAMPLE OF A SAMPLE SURVEY IN EXTENSION EVALUATION

An excellent example of the use of a statistical tool (the probability sample design) in extension evaluation study is presented in the report, "The Cass County Study," prepared by the Iowa State College Committee on extension evaluation. Although there are 9 members on the committee, my contact was with only 2 of them, Dorothy Cooke of the Statistical Laboratory and Neil Raudabaugh of the Iowa Extension Service. Also my contact in this study lasted only through the planning stage, as I left the Statistical Laboratory to come to Cornell.

The Iowa Farm and Home Labor Saving Show was placed in trucks and trailers and moved from city to city. The local extension people and chambers of commerce publicized the show. The caravan consisted of exhibits demonstrating recommended household and farm practices and arrangements and was attended by over 80,000 people in 30 different counties. Extension specialists were on hand to explain the various exhibitions.

It was not known beforehand how much the farm people of Iowa already knew about the exhibits. If the show were to be a success, the people attending the show would have to learn something new about the practices suggested and they would have to adopt some of them. Therefore, the objective of the study was to obtain valid and objective evidence of the accomplishments of the Farm and Home Labor Saving Show and the weaknesses in this method of doing extension teaching. Since a number of towns had already been visited before Mrs. Dorothy Cooke approached the Extension Service group, it was decided to try to determine the accomplishments in one area, namely the Atlantic, Cass County, trading area.

A sample survey of the selected area was made prior to the visitation of Atlantic, Iowa, by the caravan to determine the pre-show status of the practices recommended, the recognition and use of Iowa State College and Extension Service personnel and facilities, the acquaintance with county extension personnel, and the attitudes toward change.

A second survey was conducted one month after the show to determine the acceptance of the show and its influence on those attending the show.

A third survey was conducted 7 months after the second survey to evaluate the effectiveness by finding out how many intended practices had actually been adopted.

The actual sample design will not be discussed here. The reader is referred to the mimeographed report of this study. It was published in June, 1949, by the Agricultural Extension Service, Iowa State College, Ames, Iowa.

V. INTERPRETATION OF THE RESULTS OF SURVEYS

In conclusion, I would like to say a word or two on the validity of figures. American people, and perhaps other people too, are beginning to accept averages, percentages, ratios, etc. as if they were the real thing. They do not determine whether or not the figure has validity, whether or not it is possible to obtain such a value, or whether or not the interpretation of the writer is at fault. What I am trying to say is that people just aren't critical. They accept too much in the way of figures, since at this particular stage they happen to be "figure conscious." I guess that I should say "number conscious" in order to be precise.

For example, the following story appeared in Readers' Digest: The number of traffic accidents occurring at intersections having traffic lights was about 9 times higher than at intersections not having traffic lights. The reporter contacted a number of large cities and found that this ratio varied from seven to one to ten to one. He then expounded on the relative dangers of the two types of intersections. The obvious conclusion appears to be that we should do away with all stoplights. The fact that there are many times more pedestrians and automobiles at intersections with stoplights than at those without seems to have escaped the attention of the writer. This was probably intentional, as the writer was trying to bias the reader and thereby make him more conscious of the dangers.

Another fictitious use of numbers in the form of ratios is demonstrated by the advertisements of cigarette manufacturers. In particular, Lucky Strike states that the preference for their product is "2 to 1 over all other leading brands" and that these results are from a nationwide survey. First let us ask about their sampling design. We may find that two men are sent out, one to give away cigarettes and the other to interview people as to the cigarette of their preference. The "giver" stations himself a short distance from the "interviewer." The "giver" stops people and offers them a cigarette. If the person smokes the same cigarette brand, and "lights up," the "giver" walks a short distance with the recipient. The interviewer stops two people who smoke the brand to one who does not. Therefore, the summary results from a "nationwide" (probably a few towns) survey "show conclusively that it's Luckies two to one."

Also, on various Philip Morris radio programs, Bud Collier stops a person who swears that he doesn't know Collier or anybody connected with Philip Morris. Something is screwy someplace, because just by chance alone, someone should say that their brand is milder than Philip Morris. I have been told by smokers that if one reverses the process, that is, if the smoker first blows the smoke of his own brand through his nose, his own brand will seem milder than Philip Morris.

The stories are all amusing (at least I hope that they were) but my main point is that we should be critical in reading reports. We should not accept every percentage, every average, or every ratio. This critical attitude will do much to improve our own reports.

/ / / / / /