Sample Survey Design

A. L. FINKNER, Research Triangle Institute

•MANY PROBLEMS in highway transportation and safety are not amenable to controlled experimentation, and the investigator must resort to observational studies. In other words, the phenomenon under scrutiny must be examined and measured as it exists and without disturbing it. Observational studies have less investigative power than controlled experiments; although relationships can be established, it is often quite difficult to assign cause and effect. The techniques used most often to study these phenomena are commonly referred to as sample surveys.

Surveys may be classified into 2 broad general types or groups: descriptive and analytic. The investigator in the highway transportation field makes use of both types.

As it implies, the descriptive survey provides a description of the population from which the sample is drawn. A well-designed sample survey not only provides estimates of the various characteristics of the population under investigation but also should provide estimates of their precision as well. The results of descriptive surveys are often used as a basis for administrative action. They also may be exploratory in nature and used to develop hypotheses that will be tested by techniques with more investigative power.

An example of a descriptive survey is a litter composition study designed by Research Triangle Institute under the auspices of the Highway Research Board and Keep America Beautiful, Inc. The data were collected through the cooperation of the state highway departments in the various states that participated. Here the objective was to estimate the composition of highway litter so that the various industries contributing the products that ended up as litter could be made aware of the magnitude of the problem. Although the distributions by various domains were compared, the study was not designed to measure associations among the various variables.

Analytic surveys attempt to go beyond pure description and to determine relationships or test hypotheses. Here, for example, we may wish to test the efficiency of driver training or safety belts on reducing accidents or serious injuries. Although it might be ideal to be able to select a sample of young men and women of a certain age, match them with respect to certain variables, assign them to driver training or no training at random, and then measure their safety records under the same conditions over a period of years, such an ideal is unattainable. Therefore, we must resort to other techniques to obtain answers to the questions raised. Cochran (1) discusses the problems in planning observational studies and some of the current strategies in overcoming them. The problems can be classified into those encountered in setting up comparisons, dealing with disturbing variables, going from measures of association to the elucidation of causation, generalizing from the sample to the population, measuring, and considering multiple variables.

For purposes of illustration, consider 2 problems: (a) comparison of the number of serious accidents among safety belt users and nonusers, and (b) effect of various community campaigns on the incidence of drinking drivers.

In the first investigation, we must measure at least 2 quantities. We must first have an estimate of the number of serious accidents that occur to both safety belt users and nonusers. There are certain problems encountered here. The definition of a serious accident may be troublesome. In some accidents the determination will be obvious, but

Paper sponsored by Special Committee on the Conduct of Research and presented at the 49th Annual Meeting.

border-line cases may be difficult. There also may be a problem of underreporting but, in the main, these problems are tractable. The sample is usually defined in time or space or both; i. e., certain areas, such as counties or states, may be selected and examined at random times.

Before any meaningful comparisons are possible, however, an attempt must be made to match, at least on major variables. One obvious major variable is the number of drivers in the 2 categories and the number of miles driven by each, or the number of driver-miles. This is sometimes referred to as exposure and must be estimated from an entirely different type of survey. At the time exposure is estimated, other relevant variables should be measured. These include descriptions of the driver and vehicle and characteristics of the highways and the environment. Of course, these same characteristics should be enumerated for each serious accident. Even then conclusions may be tenuous. For example, if all nonusers are driving old cars that were manufactured before seat belts were required, the results could be the result of the age of the car and not the failure to use seat belts unless this kind of factor is adequately assessed. Thus, in studies of this kind, 2 surveys are required, one to estimate the numerator of a variate and one to estimate the denominator; and both are equally important.

In the second problem, let us consider some proposals being made to test the effectiveness of various countermeasures to drinking and driving. Howard Pyle, National Safety Council president, is quoted as saying that attempts to talk people out of drinking and driving have failed. The approach now seems to be to try to teach the people the amount of alcohol they can consume and still drive safely. Studies are being designed in several major cities in the United States to evaluate the effectiveness of drinking-driving countermeasures. Some countermeasures that have been proposed are (a) a program of public information and education; (b) a program of strict law enforcement in conjunction with special handling of cases by the courts; and (c) programs a and b operating simultaneously. It is hoped that the effectiveness of these countermeasures can be evaluated and that, if any show promise, they can be adopted on a larger scale.

First assume that we have only one countermeasure and one city. The so-called before-and-after type of study is often employed in situations of this kind. At a minimum, the incidence of drinking drivers must be estimated and the principal disturbing variables such as exposure by various characteristics of the driver, vehicle, highway, and environment must also be measured. At an appropriate time after the introduction of the countermeasure, the same variables are measured again. This enables us to examine whether changes in the drinking driver incidence have occurred over and above those expected from changes in the disturbing variables. The estimate of the effectiveness of the particular countermeasure is subject to the following 2 types of bias: (a) people's behavior immediately prior to the program might be affected by the knowledge that the program is about to be initiated; and (b) some disturbing variables that affect time changes may be unknown. The first type can be controlled in this case by refusing to allow any publicity prior to the introduction of the countermeasure. The second type is more serious and may be impossible to control.

Now, if it is possible to conduct the same program in 2 cities that are somewhat evenly matched with respect to disturbing variables, it is possible to select one city at random to be the programmed city and the other to be the control. Even if they cannot be matched on the disturbing variables, it is usually possible to adjust for such differences. Then the differences between the before-and-after measurements in the 2 cities are compared. Although we still have the problem of unknown disturbing variables affecting the 2 cities differently, the probability of success in measuring the effectiveness of the countermeasures is greatly enhanced by the addition of the second city.

When we have only 1 city and 3 countermeasures, like those mentioned earlier, our difficulties are compounded. Not only do we have the same problem of unknown disturbing variables, but we have an unknown contribution from a carry-over effect from the previous countermeasure when a new one is adopted. Again, these effects can be best measured by adding cities that are as alike as possible and by applying the principle of cross-over designs from the field of experimental design. The order of the "treatments" or countermeasures must be designed so as to permit estimates of the carry-over effects and then the various orders assigned at random to the cities.

The estimates of the incidence of drinking drivers also deserves some attention. Roadblocks can be set up at random points and at random times. However, the driver can refuse to be examined, introducing the possibility of a bias of nonresponse. So long as the nonresponse rates are approximately the same in each period of countermeasure activity, this bias probably can be ignored.

In summary, analytic surveys or observational studies should be designed as carefully as possible so as to minimize the bias from potential sources and at the same time to allow as much precision as possible. As near as possible, the population sampled should be the population about which inferences are to be drawn. If these inferences are to have validity, probability samples should be used wherever possible. Disturbing variables should be identified, and attempts to match or adjust for the most important of these should be made. In making the transition from measures of association to causation, the investigator will usually have evidence from a heterogeneous collection of results of varying quality. He must weight these results, giving low weight, of course, to poor quality information. Finally, he should state his judgments about conclusions clearly, attempting to be as objective as possible.

REFERENCE

1. Cochran, W. G. The Planning of Observational Studies of Human Populations. Jour. Royal Statistical Soc., 1965.