American Journal of

PUBLIC HEALTH

Official Monthly Publication of the American Public Health Association, Inc.

Volume 44 June, 1954 Number 6

On the Use of Sampling in the Field of Public Health

Here are methods with meaning to every public health worker who deals with figures—and who does not?

CONTENTS

Prepared by the Committee on Sampling Techniques in Public Health Statistics, Statistics Section, American Public Health Association, for the use of program administrators.

INTRODUCTION

✢ Sampling, in its broadest sense, is a procedure in which one or more members of a population of people or things are picked from the population. The objective of selecting this subgroup, or sample, is to make certain observations upon the members of the sample, and then, on the basis of these results, to draw conclusions about the characteristics of the entire population. When sampling is used instead of examining every member of the population it is usually for the purpose of saving money or time.

The conclusions which we draw from a sample are necessarily subject to one special kind of error, generally called sampling error, because the data have been collected for only a part of the population. However, with modern technics, a sample may be planned so that the probability is small that this error will exceed certain specified limits. When we employ sampling of this type we can say in advance how much sampling error we are willing to tolerate and then select the sample so that our statistical conclusions about the characteristics of the whole population will not have a sampling error greater than the tolerable limits, except very rarely. This type of sampling is known as probability sampling, and it is to the subject of probability sampling in public health work that this report is addressed.

The use of probability sampling in the field of public health has apparently been hampered by two widespread and somewhat contradictory attitudes toward it. The first is that sampling will not give sufficiently accurate results for a particular job at hand. The second is that probability sampling is too difficult or expensive when compared to the study of a convenient "chunk" of the population, such as a convenient batch of records, a convenient community, or a group of patients to whom the investigator has ready access.

When sampling is about to be rejected because of the belief that it will not give sufficiently accurate results, a re-examination of the errors inherent in a complete coverage may reveal that the additional error due to sampling is an insignificant part of the total error. The belief that sampling will not be accurate enough also may arise because the investigator has not stopped to work out the major objective of his investigation or to consider how the results will be used. Whether or not sampling is used in such cases, a consideration of whether it could be used brings about a clearer understanding of the purposes of the investigation and, thus, has a beneficial effect upon the conduct of the work.

The second attitude mentioned above is that probability sampling is too expensive compared to what is sometimes called "chunk" sampling. The administrator may find us quite persuasive when we say that it is often unnecessary to spend the money required for complete coverage. It is more difficult to be persuasive, however, when we say that it is sometimes necessary to adopt the more expensive of sampling alternatives. Saving money is a worthy objective. It is, in fact, one of the guiding objectives of the probability sampling we wish to describe in this report. Hence, to give this question its proper share of attention we shall put off to a later section a discussion of the relative merits of probability sampling and "chunk" sampling.

Aside from these two negative attitudes about probability sampling, undoubtedly the most serious obstacle to the greater use of this technic in public health is ignorance of the methods involved.

The Committee on Sampling Techniques in Public Health Statistics of the Statistics Section, American Public Health Association, is of the opinion that sampling can profitably be employed for many public health jobs and that beliefs of the sort mentioned previously and ignorance of sampling methods have hindered the greater use of what can be a very valuable tool. The committee has prepared this report in order to describe in nontechnical language the nature of sampling, to give a little idea of what is involved in the practice of sampling, to give a few examples of the use of sampling in public health work, and to list some references for further reading.

In writing this report we have leaned upon three sources in particular. One is "A Brief Statement on the Uses of Sampling in Censuses of Population, Agriculture, Public Health and Commerce," prepared by W. Edwards Deming, a member of the Sub-commission on Statistical Sampling of the United Nations Economic and Social Council.[1] This statement was based upon a working paper of the sub-commission bearing the reference code E/CN.3/ Sub 1/W.1. In some places we have paraphrased this document and in others we have quoted from it. We are indebted to Dr. Deming and to the sub-commission for permission to make this use of their publication. For the two other sources that have been used extensively see references 2 and 3.

THEODORE D. WOOLSEY, Biostatistician, Division of Public Health Methods, Public Health Service, Washington, D. C., Chairman
WILLIAM G. COCHRAN
DONALD MAINLAND, D.SC.
MARGARET P. MARTIN, PH.D.
FELIX E. MOORE, JR.
ROBERT E. PATTON

I. THE NATURE OF SAMPLING

What Sampling Is and Some of Its Advantages

The idea of achieving speed and economy or both by using a sample to represent the whole is so commonplace in everyday life that one is likely to overlook some of its applications. The housewife who snaps a few beans in the basket to see whether they are fresh or not is using sampling to decide whether the average quality of the beans is worth the price she is asked to pay. The great bulk of this everyday sampling is of an intuitive type and not much is lost in most cases if a wrong decision is made on the basis of the sample.

In administration and research, on the other hand, a lot may depend upon the decisions that are made. Hence, a more scientific type of sampling is called for—some method by which we may know what the limits of error are for the conclusions that we draw from the sample. We should like to be able to give some sort of guarantee about the size of the sampling error so that we can weigh the economy of sampling against the chances of error of a certain size and decide whether the savings can be realized without sacrificing too much in the way of precision. It is here that probability sampling makes its contribution.

But let us consider in more detail the returns that we may expect for introducing this additional error due to sampling. In the words of the Sub-commission on Statistical Sampling of the United Nations:

> The main advantages of sampling, in comparison with complete enumeration, are speed and low cost, with controlled and often enhanced reliability.
>
> Any survey, or any collection or tabulation of information from existing records that can be carried out completely can also be carried out by sampling. Under difficult conditions sampling is often possible when a complete count would be a failure.[1]

Certainly, the most frequent reason for deciding upon sampling is that the cost is lower than complete coverage of the population of persons or things. In fact, sometimes the cost of carrying out a test or an examination on every member of the population is prohibitively expensive. For example, a survey of the nutritional status of children in a community may involve elaborate and costly tests which can only be carried out by using small samples of the child population. In considering reduction in cost through the use of sampling, however, it is important to remember that the cost of a sample of 1/rth of the population will usually be greater than 1/rth of the cost of covering the population completely. There are certain overhead costs in sampling which are not encountered in complete coverage.

The fact that sampling cuts down the time necessary for making the information available is perhaps the next most frequent reason for using sampling. This is particularly important when the sample is being used to measure something that is changing rapidly. Again, there is some overhead in time associated with sampling, so that the information is not available in 1/rth the time from a 1/rth sample.

Sometimes in a study or survey rather elaborate information is required which is to be obtained by interview or questionnaire. In this case, a sample involves far less burden upon that part of the public from which the information has to be obtained, whether it be housewives, doctors, clinic patients, or others. The fewer who have to take time off to answer the questions, the better.

Still another advantage of sampling is that a small office can sometimes carry on the work without increasing its staff or making use of outside resources, while complete coverage might be out of the question. Even when outside resources are required for a special investigation, the problems of hiring part-time people and finding space in which they can work are much less acute when sampling is employed. In health surveys it is sometimes impossible to train a sufficient number of interviewers to conduct a detailed interview in a prescribed manner in every household of the population being surveyed. A sample offers the only feasible way of carrying out the survey. This is particularly likely to be the case if the definitions employed are necessarily difficult or if experienced judgment by the interviewer is involved, as occurs in measuring the degree of disrepair in housing surveys.

The use of sampling often permits a higher quality of work to be maintained. The intensive training of an interviewing or coding staff is more easily accomplished when the staff is small. Furthermore, higher standards of ability and personality can be maintained in selection of the staff. For example, the coding of diagnostic material during the processing of morbidity surveys is a complex operation. Small crews of coders can be more carefully selected, trained, and supervised and can, therefore, give more uniform and more accurate results.

The very nature of sampling directs attention to errors, not only those resulting from the sampling but the other types which would be present even in complete coverage. It always becomes important to get some clearer ideas of the magnitude and direction of these other errors in a sampling investigation, and the effort to do this may have a wholesome effect on the quality of the results and certainly will make for more careful interpretation of the data.

Finally, there are instances when sampling is the only reasonable possibility. Such is the case when a test is being made that destroys or makes useless the object being tested. For example, in the routine testing of bottled milk for the presence of bacteria, complete coverage is out of the question. Thus, in some instances, complete coverage of the population is not only impractical but logically absurd.

What Sampling Can Do

The following list will show how very broad are the possible applications of the sampling tool in the everyday work of public health. Sampling can be used:

To obtain health information about the persons living in a health department jurisdiction for the planning of a new program.

To test the efficacy of a proposed health department procedure before deciding whether to put it into full-scale use.

To evaluate periodically the results of a health department procedure, such as a health education program.

To survey the environment in a health department jurisdiction, as in sanitation surveys and surveys of housing.

To measure the utilization of health services or the availability of medical care in the jurisdiction.

To obtain information from a large file at less expense and more quickly than would be possible by complete analysis.

To prepare preliminary bulletins on vital statistics in advance of final processing of the records.

To evaluate the reliability or completeness of registration of births and deaths or other record systems.

Of course, there are many other uses. No division of the health department, no voluntary health agency, and no health research institute is without its own applications. It has been demonstrated many times that the application of sampling can bring about savings in routine office work as well as in special studies and surveys. Sometimes it is a matter of applying sampling where it had not been realized that it could produce sufficiently reliable results. At other times methods can be suggested for controlling quality at less expense. The use of sampling for quality control has found widespread favor in industry and the application has also been made in such large-scale, government statistical operations as the processing of records from the decennial census and those of the National Office of Vital Statistics.

Two continuing sample operations in the field of public health and medical records that may be mentioned are: (1) the sampling of sickness records for Army personnel; and (2) the 10 per cent sample of death records analyzed each month by the National Office of Vital Statistics.

Reliability of Sampling

How reliable is sampling? The answer is: Sampling can be just as reliable as we need to make it and can afford to make it. The decision that nothing short of complete coverage of the population will give a sufficiently reliable statistical picture is often made without real justification, probably because of failure to consider two important aspects of reliability. The first is that the error we can tolerate in statistics is related to the purpose the statistics are to serve. When that purpose is not clearly recognized in advance there is quite understandably a tendency to seek refuge in complete coverage. More attention to the matter of final use of the statistics will frequently show that a sample can give sufficient accuracy for all major uses and, when this discovery is made, the additional time spent in planning is more than paid for by the lower cost of the investigation itself.

The second aspect of reliability to which attention should be directed is closely related to the first. It is the realization that sampling error (which, it will be remembered, is the error resulting from taking a sample rather than the whole population) is only one of the possible sources of error in the results. Nonsampling error may be of many different types and the types vary according to the kind of investigation being made. In the case of one kind of investigation—the survey—the different sources of error have been catalogued in a very interesting and useful chapter of a book by Dr. Deming.[2] Dr. Deming lists 19 different sources of procedural bias and error. We cite some of these to give an idea of their nature. Most of them are sources of error with which everyone is familiar, but familiarity seems to have led to excessive tolerance in many cases. Here are 10 different sources of error that are found in surveys:

1. Failure to state the problem carefully and to decide just what statistical information is needed.
2. Failure to define the population about which we want information with sufficient precision.
3. Errors in response, voluntary and involuntary, i.e., errors in the answer to a question due to misunderstanding, misinformation, misrepresentation, faulty memory, etc.
4. Bias in response arising from the interviewer.
5. Imperfections in the design of the questionnaire and tabulations of the results.
6. Bias arising from nonresponse, i.e., bias from the fact that answers were not obtained for all in the sample and from the fact that these nonrespondents differ from the respondents in the characteristics being measured.
7. Bias arising from an unrepresentative date for the survey or of the period covered.
8. Random sampling errors. (See below.)

9. Sampling biases. (See below.)
10. Processing errors (coding, editing, calculating, tabulating, tallying, posting, and consolidating).

Of the 19 sources of error listed by Dr. Deming only two are found in sampling but not in complete coverage of the population. The author calls these two "random sampling errors" and "sampling biases." The first is the kind of error with which probability theory deals and the second arises either from human failure in carrying out the sampling instructions or from technical failure in the process of generalizing from the sample to the whole population.

Dr. Deming has this to say of sampling error:

> One often hears objections to sampling because of sampling errors. Such objections can be sustained only if, after consideration of the other inaccuracies, the elimination or reduction of the sampling errors seems to be a wise investment. Sampling errors have the favorable characteristics of being controllable through the size and design of the sample. . . . Sampling errors, even for small samples, are often the least of the errors present. (See reference 2, p. 47.)

In its general outline this consideration of sampling and nonsampling errors is not peculiar to surveys, but is common to any problem of measurement, whether it be the taking of blood pressures or the decennial census.

But what do we mean when we say that sampling errors are controllable? We obviously do not mean that we can tell how great the error is in a particular sample and in measuring a particular characteristic. What we do mean is that we can give a very useful kind of guarantee regarding the sampling error. Let us take as an example a hypothetical investigation of the children in the elementary school system of a city to determine the proportion of children with vision below a certain standard. Vision below this standard we shall call "impaired vision." We have selected completely at random a sample of 400 children from the elementary grades of the school system and have determined by means of tests that 4 per cent of the children in the sample have "impaired vision." In this particular investigation the guarantee might take the following approximate form: "The chances are 19 out of 20 that the 4 per cent of children with impaired vision, as estimated from the sample, does not differ by more than 1.9 per cent from the result that would have been obtained if the identical test had been given to every child in the elementary school system."

The range of approximately 1.9 per cent on either side of the estimated 4 per cent is known as the "confidence interval" of the estimate. The principles of sampling tell us that if we were to repeat this investigation over and over again, each time making an estimate from a new sample of 400 children in an identical manner, and were to compute confidence intervals as has been done here from the results of each sample, in 19 out of 20 instances the confidence interval would include the city-wide value, i.e., the value which would be obtained if the test had been performed upon every child in the elementary school system.

In the planning of a sample we wish to know the most economical type of sampling plan and to estimate the size of sample required to give the desired precision. For both of these purposes we must have some advance knowledge of the size of sampling error associated with samples of different sizes and designs. Sampling principles can provide us with this equally useful but less specific type of information in advance of the sampling, providing that we know roughly what the results will look like (as we frequently do by examining results obtained at other times or in other places).

The reason some foreknowledge of

the approximate results is needed is that the formulas for the sampling error associated with different sampling plans require information on two points: the mean value of the measure (or the proportion of the population having the characteristic being studied) and the amount of variability in the measure from member to member in the population, expressed in terms of the standard deviation. For example, if we were planning to survey a sample of houses in a city to estimate the average state-of-repair score of houses in the city, we would need to know roughly what the average score and the standard deviation of scores from house to house would look like. The better the information we have on these points, the more accurately we can estimate the sampling error of the survey results in advance, but even informed guesses will be helpful. (If the result desired from the survey is simply the proportion of the population having some characteristic, a knowledge of what that proportion is likely to be is sometimes enough.)

Note that these statements concerning limits of error provide no guarantee on the nonsampling error. Nothing is said, for example, about the error that may be implicit in the test itself. No test is perfect and the one given in the vision test illustration mentioned previously may be one which, if repeated upon the same child, would show a great deal of variability in results, either because of reading errors by the person making the test or because of some lack of precision in the measuring device. Such errors implicit in the test will be present in the results of an investigation even if it covers the entire school population. Nevertheless, the use of the principles of sampling both for planning the sample and for a statement of the limits of sampling error in the final result are the features of probability sampling that make it possible for us to use it with confidence.

The sampling error that we feel we can tolerate and the chances that we are willing to accept of being off by this amount can be altered to suit our convenience, but, in general, the smaller we demand the sampling error shall be, the more expensive the task of measurement becomes. This is because reducing the sampling error means increasing the size of the sample and this, of course, increases the cost. (Improvement can sometimes be made without increasing cost by devising a more efficient plan, but there is a point beyond which one cannot go in improving the sample design. The only way to reduce sampling error from this point on is to increase the size of the sample.)

The use of sampling principles in advance of the actual collection of data is valuable to the statistician and the program administrator alike. In the case of an expensive survey it might work something like this: The statistician first gathers whatever information is available from other sources as to how much the particular characteristic being measured or counted varies from one unit of the population to another. He also brings together whatever facts are available about the cost of the different operations involved in collecting the data. If there is no information on hand regarding these matters, he may advocate the conducting of a small pilot study to collect it in advance of the full-scale project. Such a pilot study would also provide an opportunity to test the measuring device itself, that is, the interview, the clinical test, the examination procedure, or whatever else was to be employed for the purpose of measurement or description. With the information on variability and costs and a knowledge of sampling technics, the statistician can determine which of many possible plans will probably give the least sampling error for a fixed sum of money; or, for a required precision, which plan will cost the least.

In our hypothetical survey of impaired vision, for example, the results of other studies of impaired vision among school children might permit the statistician to devise a probability sampling plan that, without sacrificing any of the sampling precision, would cost considerably less than a selection of 400 children completely at random. The statistician turns over to the administrator his estimates of cost and expected limits of sampling error. Thus, the administrator can tell fairly reliably before embarking upon the project whether the amount he is able to spend will yield results of useful precision, though in making his decision he must also consider the errors from sources other than sampling (see below).

This, then, is what we mean when we say that sampling errors are controllable through the size and design of the sample: By skillful use of existing resources and knowledge a plan of sampling can be devised such that the sampling error is reduced to the minimum for a given expenditure of money; and, furthermore, the limits of sampling error, or "confidence interval," of the estimate made from the sample can be stated alongside the estimate itself.

But what about the many types of nonsampling error, some of which we have cited from Dr. Deming's list? Can these be controlled too? Such errors can be controlled, but not in the same sense that sampling errors can be controlled. Some nonsampling error is always present and increasing the size of the sample will not reduce it. It can be reduced, however, by careful planning, including the making of trial runs. Even if the sampling consists of no more than the selection of every tenth card from a file, when any large amount of work is involved a trial run will pay off. It will show up unexpected difficulties before it is too late to make changes in the plans.

Research on the magnitude of the errors implicit in procedures of measurement, such as the errors mentioned earlier in connection with testing of vision of school children, should not be overlooked. If the investigator has no quantitative information on this point, he will do well to make the study of errors of measurement a part of his investigation. As one example of this type of research the reader is referred to the study of the errors in counting blood cells made some years ago at the Mayo Clinic. The opening sentence of the report on this study reads: "Our primary objective was to investigate the error of the blood cell count as it is made in good routine practice, and not as it might be made with special apparatus in exceptional situations."[6] Another example, unfortunately unpublished, is the study of the errors in blood pressure readings with a sphygmomanometer made some years ago at the Worcester State Hospital in Massachusetts, by means of a special stethoscope with multiple headpieces designed so that simultaneous readings of the same blood pressure could be made by several doctors.

There is not space here to cite the considerable experience that has been accumulated on the matter of reducing nonsampling errors.* However, it must be emphasized that it is this part of the planning of a survey, or other investigation, that frequently requires the greatest care and attention. This fact has a bearing on the answer to the perennial question: "If we are to employ an expert to take charge of the collection of data for the survey, what sort of expert should we look for?" The person to look for is someone who has had enough experience with surveys to be able to keep the nonsampling errors at a minimum. If not an expert in the particular subject matter that is to be dealt with, he must be supported by a

---

* References 2, 4, and 5 will be helpful in this respect.

technician thoroughly familiar with the subject matter.

It has already been mentioned that the control of nonsampling errors is usually easier in sampling than when coverage is complete. This is because there are fewer workers to supervise and because more attention can be devoted to detail. If interviewers or other trained personnel are being used to collect the data, there are fewer to train, and greater uniformity and intensity of training is possible. For instance, 20 persons might have to be taught to count the number of decayed, missing, or filled teeth among the patients of a large institution. A 25 per cent sample would require the training of only five. It is conceivable, therefore, that an estimate for the whole institution based on a 25 per cent sample could give a result closer to the true value than one obtained from an examination of every patient.

We have said that the program administrator must consider both sampling and nonsampling errors in considering how much sampling precision to buy. To illustrate, let us say that he believes he cannot tolerate a relative error due to sampling of more than 2 per cent. It may become apparent on further study, however, that because of some intrinsic error in the method of measurement (such as, for example, a lack of precision in the test of vision in the example cited previously) there is a possible error of as much as 10 per cent in the result even if every member of the population is included in the investigation. In that case the administrator has the choice of dropping the idea of making the study, of revising upward his ideas of tolerable sampling error, or of spending money to attempt to reduce the nonsampling error by some means. This last could be done by devising a more precise test or, in other types of studies, by adopting more objective definitions or clearer instructions, or by intensifying the program of interviewer training, etc. A good deal of attention has been given here to this matter of the reliability of sampling because it is the cause of so much misunderstanding. Furthermore, the heart of probability sampling is the controllability of sampling error.

Disadvantages of Sampling and Situations in Which Sampling Will Not Help

It is important to bear in mind that sampling is not an end in itself. It should not be thought of as a procedure that adds elegance to any job that comes along. Its use is suited to certain situations and unsuited to others, and it has its disadvantages. Before proceeding with a discussion of the practice of sampling it might be well to mention some of these disadvantages of sampling and situations in which it cannot help.

Where an Inventory Is Needed—Obviously, sampling cannot help where a record is needed for every member of the population of persons or things. A hospital could never make medical history sheets for only a sample of its patients because actions affecting an individual patient must be based upon a report concerning that individual alone. The same would be true of a case register. However, if a hospital wishes to draw some general conclusions about a certain class of its patients, a sampling of the medical records for that class of patients may give the desired results at great savings of money and effort.

Data Required for Many Subdivisions of the Population—There are also occasions when data must be tabulated in great detail in order to be valuable. Sampling would make the frequencies in some cells of the tables so small as to be useless for purposes of analysis. In the decennial test of completeness of birth registration, for example, census enumerators fill out a card for any infant less than three months of age living in the household at the time of the census. The birth registration records are searched for the names on these cards and an index of completeness of registration is computed. A good deal of money would be saved if this could be done on a sample basis. However, incompleteness of registration has become a highly localized problem. Action to improve the registration depends upon knowing the particular county in which it is poor. Hence, a determination of the size of sample required led to the conclusion that all infants less than three months old would have to be included.

Additional Error Introduced—In sampling, as we have said, an additional source of error is introduced. While the limits of sampling error can be calculated approximately in advance and it can in this sense be controlled, there still may be occasions when nothing less than a complete coverage of the population will be sufficient. Such occasions are not as common as one might think. One instance that can be cited is the reporting of cases of an epidemic disease. Health department action may be based on the reporting of a very few cases, and outbreaks of many cases may be missed in a sample because of their concentration in space and time.

Time and Skill Required in Planning—Since the planning of a good sample may require more time or technical skills than are readily available, this lack may sometimes argue against sampling. However, as has been pointed out, the planning time may be well spent because it gives a greater opportunity for the objectives to be clearly thought out and for misunderstandings to come to light.

Discrimination Against Those Not in the Sample—A peculiar situation may arise when a new treatment is being tried on a sampling basis or when sampled individuals are given some special and desirable service, such as a free physical examination, an x-ray, or free treatment if a case of illness is discovered. This has been handled in at least one study by providing the same service to all who heard about the offer and asked for the service during the course of the sampling. This might well be impractical in some projects, but in these projects it stands to reason that complete coverage would also be impractical.

II. SOME REMARKS ON THE PRACTICE OF SAMPLING

When to Call for Expert Assistance

We have said that there are jobs of sampling that can be carried out in a health department or other public health agency without the help of a sampling expert. We have also suggested that there are times when it will be wise and economical to call for outside help. Assuming it has been found desirable to sample, how does one decide whether to seek expert assistance or not? There is no hard and fast answer to this question. A lot will depend upon the amount of experience with sampling that is to be found in the staff of the office. However, the following matters should be looked into before making the decision:

The Cost of the Investigation—Is the study one in which a considerable sum of money is to be spent? If so, it is undoubtedly necessary to adopt a plan that will give the utmost in information for every dollar spent. The cost of obtaining the advice of a sampling statistician will probably be a small portion of the total and in many cases there will be a net saving.

The Complexity of the Task of Drawing a Satisfactory Sample—To attempt to grade sampling studies according to the complexity of the sample design required would take us further into the theoretical aspects of this subject than we intended to go in this report. There are, however, some clues to the probable complexity of the problem of sample selection that are supplied by a knowledge of the population being investigated and the resources at hand.

Sampling implies the physical selection of certain members of the total population. There must, therefore, be some catalogue or inventory of the population from which to draw the sample. If no such inventory exists, it must be constructed. For example, there is the common type of epidemiological study in which a sample of cases of a particular disease is compared with a set of matched controls to determine whether the cases differ from the controls in respect to some characteristic suspected of being related to the risk of contracting the disease. In these studies it is often necessary to create a register, or list, of cases of the disease in a specific population. Making up the register may require searching the records of doctors' offices, clinics, and hospitals of an area for all cases of the particular type being investigated. A list of persons to be used as controls may also have to be constructed. When these lists have been prepared a sample of persons with the disease can be selected and matched with persons in the list of controls.

The first question one might ask, therefore, in sizing up the problem of selecting a sample is: Can the population of persons or things that is to be sampled be represented by a file or list, each item of which represents a single member of the population? If so, is this file or list complete? Does every member of the population that is under study appear in the file or list once and only once? If the answer to these questions is "Yes," the sampling can usually be accomplished very simply. Very little additional difficulty is encountered if the file or list contains some items that are not members of the population we wish to study, but if there are duplicates or if some members of the population are not represented at all, the difficulties become greater. (These conditions do not necessarily preclude sampling; but expert assistance should probably be sought.)

Sometimes there is in existence a list or file that includes the entire population but the individual cards, folders, or list items are not the units we intended to sample. For example, we might set out to select a sample of families but find that the file contains a card for every adult person rather than a single card for each family. Or we might need a sample of individuals and have only a file containing a card for each family with the names of the members listed on it. In these cases in which the unit member of the list or file is not the same as the unit member of the population we wish to study there is some additional complexity.

If no list or file exists in which every member of the population is represented, we must consider whether one can be constructed, or whether the members fall in some natural order, as in the case of assembly line operations, so that a simple consecutive numbering system can be assigned to them. The problems of selecting a good sample are less difficult when the items to be sampled can be easily numbered. Two familiar examples of sampling where the units cannot be simply arranged in order and numbered are the making of differential blood counts and water sampling. The problems of devising acceptable probability samples in these cases are exceptionally difficult. The sampling of insect populations is another example of the same type.

The selection of a representative sample of the people living in a city or county or other health department jurisdiction is a common problem in public health work. Finding an up-to-date and complete directory of the families residing in a community is a rare occurrence in this country. Nevertheless, experience indicates that directories, unless hopelessly out of date, can be helpful in reducing the cost of sampling, but this comes under the heading of problems for which expert assistance is desirable.

Very often there is no directory at all from which to sample. Usually the method employed in such cases is that known as area sampling. The inventory is constructed from a map by dividing it into small units of area. A sample of these small areas is then drawn. This is, therefore, an instance of the type mentioned earlier in which the unit actually sampled is not the same as the unit member of the population we wish to study. Area sampling problems have a great range of difficulty, depending upon the resources available and the efficiency demanded of the sample design, that is, the extent to which the design will provide the minimum sampling error per dollar expended.

The Complexity of Nonsampling Aspects of the Study—It has already been pointed out that the reduction of nonsampling errors is the part of the survey or other investigation to which the greatest care and attention may have to be given. We must ask questions about the sort of errors that would be encountered even if the study were designed to cover every member of the population. Has it been possible to state clearly the problem to be investigated? Has it been possible to define the population precisely? Is it going to be possible to reach all the members of the population without difficulty? Will people agree to answer these questions or to take this test? Are the questions of a factual and noncontroversial nature? Is the test to be given one that will yield consistent results when repeated by the same technician? When repeated by different technicians? If the answer to any of these questions is "No," it may be time to call in an expert. No miracles should be expected, of course, but sometimes a wider experience in dealing with similar obstacles will suggest ways of surmounting them.

For one class of investigation, the mail questionnaire, a special warning should be given. It seems to be very easy for those who have not had experience with this sort of inquiry to overestimate the amount of response to be expected and to underestimate the amount of bias that may be introduced by neglecting the people who do not return the questionnaires. In questionnaires addressed to a general population group the response rate seldom exceeds 50 per cent. When the respondents are members of a special group, such as doctors, nurses, directors of nursing homes, and so forth, the rate of response may vary greatly. It is probably widely recognized that the response depends upon the auspices of the study, the length of the questionnaire, the format of the questionnaire, and other factors. However, it is not so well known that the characteristics of those who do not respond may differ widely from those who do. Consequently, reliance upon statistical results of the tabulation of questionnaires from those who took the trouble to reply may lead to false conclusions. The respondents cannot in general be assumed to be a representative sample of those to whom questionnaires were mailed. How to overcome this bias and at the same time make use of the undoubted economies that can often be achieved by a mail questionnaire is a matter on which an experienced statistician's advice should be sought.

It will be noted that these comments also apply to any study or survey where the subjects are volunteers. It is generally conceded, for example, that a chest x-ray screening in which twothirds of the adult population of a community voluntarily submit to an x-ray cannot tell us very much about the proportion of significant films that would be found if the entire adult population had been screened. The proportion of positive films in the remaining one-third may be quite different.

Further Remarks on Calling in an Expert

The objection that is sometimes heard to the employment of a sampling statistician is that he will be unfamiliar with the practical aspects of a job, will confuse the issue with mathematical abstractions, and will recommend something more elaborate and expensive than anyone had in mind. With a proper understanding of the statistician's function on the part of both the administrator and the statistician such difficulties should not arise. For sampling is not an esoteric science. Its objectives are exceedingly practical. A statistician who understands good sampling is aware that no sampling plan is of value to an administrator unless it can be carried out in the field.

Furthermore, since modern sampling theory dealing with surveys and experiment design is much concerned with dollars and cents, it is incumbent upon the statistician to demonstrate that his plan will provide the accuracy required at a cost (including overhead) that is lower than any other practical plan meeting the same criteria for accuracy. Again, it must be emphasized that the sampling statistician can do a better job in this respect if he is brought in before plans have crystallized.

Brief Description of the Procedures in Sampling

It has been pointed out that when the type of sampling known as probability sampling is used we can say in advance how much sampling error we are willing to tolerate and can then select the sample so that statistical conclusions about the characteristics of the entire population will not, except rarely, have a sampling error greater than the tolerable limits. We shall now describe briefly what is involved in selecting a sample of this sort and how the confidence interval is obtained in a simple example.

The term "probability sampling" is applied only to samples selected in such a way that every member of the population has a known probability of being chosen and such that the probability is not zero for any member of the population. This is because it is only for samples selected with known probabilities that we can give the sort of guarantee mentioned in the section on Reliability of Sampling. When the probabilities are known and are not zero, sampling theory provides formulas that give the standard error of the measure being estimated from the sample, that is, the standard error of estimate.

This can be illustrated using the simplest type of probability sampling known as simple random sampling. In this type the probability that a particular member of the population will be selected is equal for all members. Thus, if we had a file consisting of cards numbered from 1 to 900 and wished to choose a simple random sample of 100 cards, we might pick 100 different numbers between 1 and 900, inclusive, completely at random. The cards bearing the selected numbers would constitute the sample. If the selection is truly random, no number has any greater or less chance of coming into the sample than any other number.

But how would one go about picking the hundred different numbers completely at random? One way of doing this would be to put 900 numbered slips of paper into a box, mix them thoroughly, and then withdraw 100 of them, recording the numbers of all that are selected. A simpler and more foolproof method, however, is to make use of tables of random numbers. Opening to any page of such tables 100 three-digit numbers between 001 and 900 are picked out in some systematic manner, such as the first 100 in the odd-numbered columns. In case a number turns up a second time in the drawing it is discarded and the next number is taken in its place. This, of course, is necessary because we want to have 100 different numbers in the sample. It is called sampling without replacement. (In sampling with replacement each time a numbered slip was drawn from the box its number would be recorded; then the slip would be returned to the box and the slips would again be mixed before another slip was drawn. In sampling with replacement it is obvious that a given number can appear more than once in the sample; hence, sampling with replacement is much less frequently used in practical sampling work.)

Now let us suppose that the 900 cards in the file from which we are sampling have on them some measure, such as the age at death of a certain type of cancer case. The sample is to be used to obtain a quick estimate of the average age at death. The formula for the standard error of estimate of the mean in simple random sampling without replacement is:

$$\sigma_m = \frac{\sigma}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

in which n and N are the sizes of the sample and the population respectively, and σ is the standard deviation of the ages at death of these cancer cases.

The statistic, σ_m, is a measure of the scatter that sampling theory tells us we can expect in the means of samples. Hence, it is a measure of the error that can be expected owing to having taken a sample rather than the entire population. It grows smaller as n grows larger, and it is easily seen that if n is equal to N (which occurs when the sample includes every member of the population), σ_m becomes equal to zero; that is, we have no error due to sampling.

The nature of the statistic σ_m is such that if we find the average age at death of the cancer cases in the sample, which may be called m, we can obtain the confidence interval for this estimate by adding and subtracting from m a certain multiple of σ_m.* The multiple used depends upon the chance that we are willing to take of being off by a given amount. For example, it will be remembered that the statement of the guarantee in the example, cited earlier, of a sample survey of school children with impaired vision, started off as follows: "The chances are 19 out of 20 that the 4 per cent . . . as estimated from the sample, does not differ by more than 1.9 per cent . . ."

Another way of stating the same guarantee would be to say: "The range, 4 per cent minus 1.9 per cent to 4 per cent plus 1.9 per cent (i.e., the range 2.1–5.9), is the range which in 19 samples out of 20 will include the value that would have been obtained if the identical test had been given to every child in the elementary school system."

When the chance we are willing to accept of being off by more than the stated amount (1.9 per cent in that example) is one in 20, we speak of the confidence interval as a 95 per cent confidence interval. For this confidence interval the multiple used with σ_m is 1.96. For a 99 per cent confidence interval the multiple used with σ_m is

* Sometimes obtaining the confidence interval is not this simple. For example, the confidence interval may not be symmetrical about the observed value of the mean. However, this is not a textbook and for more rigorous statements of theory the reader is referred to the textbooks listed at the end of the report.

samples out of 100 will include the value that would be obtained if coverage had been complete, the multiple used is 2.58. The multiples to be used with other confidence intervals can be found in statistical tables.

Returning now to the example of the card file of cancer cases, we can say that the chances are 19 out of 20 that the mean age at death as estimated from the sample does not differ by more than $(1.96 \times \sigma_m)$ from the mean age we would have obtained had we used every card in the file.

In the formula for σ_m given above, σ refers to the standard deviation of the ages at death of all the cancer cases in the file. This we usually would not know, but it, too, can be estimated from the sample. So let us suppose that we estimated the mean age at death from the sample of 100 cards as 47.8 years and that we estimated the standard deviation of ages at death as 7.5 years. Putting $\sigma = 7.5$, $n = 100$, and $N = 900$ into the formula, we come out with

$$\sigma_m = .707$$

$$1.96\ \sigma_m = 1.38572 \text{ or about } 1.4.$$

Thus, from the sample of 100 cards we estimated the average age at death as 47.8 years and we feel pretty sure that, had we computed the average from all 900 cards, we would not have differed from the sample estimate by more than 1.4 years.

In this example we have started with the assumption that a sample of 100 cards was to be taken, but ordinarily we might want to use sampling theory in advance of the sampling to tell us how much error to expect from a sample of a given size or, more frequently, to tell us how big a sample would be needed to achieve a certain degree of sampling precision. For this, the same formula can be used by turning it around so that it gives n in terms of σ_m, σ, and N. A convenient way of writing the formula for finding the required size of sample turns out to be:

$$n = \frac{t^2\ \sigma^2}{e^2 + \dfrac{t^2\ \sigma^2}{N}}$$

in which e is the amount of sampling error we are willing to accept and t is the multiple of σ_m (1.96, or 2.58 or some other value from the tables) corresponding to the 95 per cent confidence interval, the 99 per cent confidence interval, or some other confidence interval. Supposing, for example, that we needed to know the average age at death of cases in the file to within one year of age, using a 95 per cent confidence interval. Then $e^2 = (1)^2$ and $t^2 = (1.96)^2$. Putting these into the formula with $\sigma^2 = (7.5)^2$ and $N = 900$, the result is found to be: $n = 174.25$ or about 174.

(Note that the sample would have to be larger than 100 because we are demanding greater precision.)

However, it will be observed that in finding the required size of sample, we made use of σ, a quantity which was to have been estimated from the sample. Obviously, we cannot estimate it from the sample in advance of sampling. In fact, in order to estimate the size of sample required we must have in advance some knowledge of the variability of the measure in which we are interested. In this example the knowledge might come from another study of a similar type of cancer, from vital statistics reports, or, if worst came to worst, from an informed guess on the part of an expert in this subject matter. It is because of this need for some advance estimate, or guess, as to the variability of the measure, that we have described this type of information provided by sampling theory in advance of sampling as "equally useful but less specific."

The basic principles that have been reviewed here can be extended to more complicated types of sampling. By way of illustration we give only a very general description of two of these.

Stratified sampling, a common variation of simple random sampling introduced in order to obtain more sampling precision for less money, is sampling in which the population is first divided up into more or less homogeneous subgroups, or strata. If the number of members of the population in each of the strata is known, simple random sampling can be done independently in each of the strata and an estimate of the measure desired, such as the mean or the proportion having some particular characteristic, can be obtained for the entire population by making a weighted average of the estimates from the different strata.

For substantial gains in precision from the use of stratified random sampling, both of the following must hold true: (1) it must be possible to divide the population into strata such that the mean of the measure that we are investigating, or the proportion having the particular characteristic, differs widely from one stratum to another; and (2) the number of members of the population in each stratum must be accurately known. Furthermore, the gains will not be realized and, in fact, the precision may actually be lowered (in comparison with a simple random sample of the same size) if the sample is improperly allocated among the strata. The simplest device for allocating the sample in stratified random sampling is known as proportional stratified sampling. This means that the same fraction of the stratum population is taken into the sample in each stratum. Even better methods of allocating the sample may be found by taking into account differences in the amount of variability and in the cost of collecting the information from one stratum to another. These optimum methods, their limitations, and the situations in which they are likely to prove valuable are beyond the scope of this report.

Two other points should be made, however, before leaving this subject. First, stratification that may be good for measuring one characteristic may be poor for another. Hence, when a sample is designed to collect information on a number of points, the sampling statistician usually sticks to rather a simple and basic type of stratification. Second, stratification may be desirable for reasons other than that of obtaining a more precise estimate for the entire population. Sometimes there may be interest in having separate estimates for each stratum, or there may be a different problem of data collection in one or more strata.

Sampling of clusters or aggregates is of particular value in surveys of human populations to save money in reaching people where they live. In this type the members of the population that is sampled are not the units in which we are ultimately interested but are clusters or aggregates of those units. For example, in sampling the population of a city a sample of blocks may be selected and all the households in the sample blocks may then be included in the sample. It is apparent that such a sample involves less traveling than a simple random sample of households of the same size. However, a simple random sample of households has less sampling error for most items than a simple random sample of blocks containing the same number of households; and, in general, sampling of clusters or aggregates will almost always be found to sacrifice something in precision when compared with samples of smaller units, the sample size in terms of the smaller units being the same. Thus, supposing a city has 20 elementary schools. A sample of five schools containing 2,400 pupils in all can be expected to be less precise for characteristics of the pupils than a random sample of 60 classrooms containing the same total number of pupils. The latter will, in turn, be less precise than a random selection of 2,400 pupils.

It is not necessary, though, to enumerate completely the entire cluster in sampling of this sort. The members of the clusters or aggregates selected in the sample can themselves be sampled. This is known as two-stage sampling. (In large-scale surveys the process is sometimes extended to three or four stages.) In a population survey of a city, for example, a simple random sample or a stratified random sample of blocks is chosen. A list of all separate households in the sampled blocks is then made and from this list a sample of households is obtained.

It may be seen that in a sample involving stratification and more than one stage, there is a great variety of ways that the total sample can be divided up. Such questions arise as: What proportion of the total sample should be taken from each stratum? How big a cluster should be taken as the unit for the first stage of sampling? Should we take more first-stage sampling units and fewer second-stage units within each first-stage unit, or fewer of the former and more of the latter? These decisions influence both the cost and the eventual sampling error of the results. It is in the selection of the combination that will give the most precision for a given sum of money, or the lowest cost for a required degree of sampling precision, that modern theory of sampling surveys affords a guide.

Three Illustrations of Sampling in Practice

In a previous section we have attempted to give some advice on deciding when to call upon expert assistance. Here we present three examples selected to illustrate sampling jobs for which expert assistance should not usually be needed. They are not to be taken as model procedures, but as representative of the many instances where a satisfactory sampling job can be done without resort to complicated sample designs.

1. Sampling from a File—The New York State Health Department had the problem of determining what proportion of the birth certificates arriving at the Office of Vital Statistics in Albany were on a new birth certificate form that had been introduced six months before. Supplies of the old form had not been withdrawn, so it was known that until these were exhausted the forms arriving would be a mixture of the new and the old. The certificates for births recorded in a particular month are numbered and are kept in numerical order.

In January, 1952, there were 12,910 births recorded. It was decided that 2 per cent of these, or 258 certificates, would provide a sample of sufficient size for the accuracy desired. A number between one and 50 was selected at random. This number was 27, let us say. Certificate numbers 27, 77, 127, 177, and so forth, were examined to determine whether they were on the new or old form. It was found that 80.6 per cent of the sample were new certificate forms. This percentage had a 95 per cent confidence interval of 75.6–85.6 per cent which was sufficiently accurate for the purposes at hand. The entire job required about an hour, while examination of all certificates would have taken at least a full day.[9]

2. Sampling from a File of Families—Sometimes files are kept by families, the information for all members of a family being recorded on a single card. In this event, the method in the previous section can be used if data about the whole family are wanted. If, however, a sample of say 400 individuals is desired, it may be advisable to have only

one person from a given family. In this event, the selection of the sample is not quite so simple. It is not satisfactory to select 400 cards (i.e., 400 families) and then pick one person at random from each family. This type of sample has an overrepresentation of the smaller families, because a person who belongs to a family of size one has 10 times as large a chance of being selected as a person who belongs to a family of size 10.

If a list of the number of persons per family is available, one method is to add these numbers in succession, making a cumulative list of numbers of persons. A sample of persons is then drawn from this list by the method outlined in the previous section. Thus, the sample might consist of individuals whose members are 35, 135, 235, and so on in the cumulative list. We must then locate these individuals on the family cards to obtain the necessary information about them.

This method is laborious if the population is large and is not feasible unless a list of the number of persons per family is already at hand. An alternative method is to construct this list from a sample of families. Suppose, for illustration, that the average number of persons per family is about four and that the population contains about 12,000 cards, i.e., about 48,000 individuals. A sample of one in 100 (about 480 individuals) is to be drawn. Inspection of some cards suggests that relatively few families have more than nine persons. Since 9/4 is 2.25, we first draw a sample consisting of 480 × 2.25, or 1,080 cards. (The reason for choosing this size of sample will be explained presently.) Since 12,000/1,080 is approximately 11, this sample can be drawn by taking every 11th card by the method given in the preceding section.

Now we form the cumulative list of numbers of persons in this sample. Since this sample will contain about 4 × 1,080 or 4,320 persons, we obtain the desired sample of 480 persons by taking every ninth person in the cumulative list. It is clear that no family with nine or fewer members will have more than one member in this sample. If families with more than nine persons are present, some of these will have more than one member in the final sample. This should happen only rarely; in order to avoid it completely, we would have to make the first sample of cards larger.

3. Sampling of Institutions—A common problem is to sample a group of institutions that vary in size. Although the best method of sampling varies from case to case, the general principles can be illustrated from recommendations made by the Commission on Chronic Illness for sampling the proprietary nursing homes in a state so as to obtain information about type of illness, care provided, length of stay, etc.[10] It is assumed that each institution must be visited in order to obtain the information, so that the purpose in sampling is to cut down the travel and time involved in these visits. There may, of course, be other reasons for preferring a complete census.

The steps in planning such a sample might be outlined as follows:

1. Be sure that the definition of a proprietary nursing home is clearly understood.
2. Obtain a list of all proprietary nursing homes that meet the definition. Some effort may be required to insure that this list is complete.
3. Usually there will be a number of large institutions that should automatically be included in the sample. Make a separate list of these institutions.
4. Group the remaining institutions into strata in order to make the sample more representative with respect to factors that are considered important. The best factors depend on the kind of data that are being sought. In the present example, rural-urban location, highest level of nursing staff, and bed capacity are suggested.

The remaining institutions are grouped into two or three classes representing rural-urban location. Within each class, arrange the institutions in two or three subgroups (strata) according to the highest level of nursing staff. Finally, arrange the institutions within each subgroup in order of some measure of bed capacity.

5. Draw a sample from each subgroup (stratum) by the method outlined in the first section, taking, say, every third institution in any stratum, depending on the size of sample desired. Since institutions have been arranged in order of bed capacity within each stratum, the selection of every third institution will give a sample in which institutions of different sizes are represented in about the same proportion. The different strata are also represented in the same proportions, so that the sample has a good distribution on all three factors introduced into the stratification.

6. In this kind of study it is often advisable to obtain data about all patients in each institution that falls in the sample. If this is done, some care is necessary in combining data from the large institutions with those from the remaining institutions, because the large institutions were deliberately over-represented in the sample (step 3). A convenient procedure for avoiding this complication, making the sample self-weighting, is to record data for every third patient in the large institutions.

"Chunk" Sampling vs. Probability Sampling

The word "chunk" is used by some statisticians to denote a sample the selection of which is dictated primarily by convenience or by administrative considerations and secondarily because, in the judgment of the investigators, it appears to be typical of the population about which information is needed. "Chunk" sampling is not probability sampling as we have described it in the Introduction. Sometimes it may have the appearance of probability sampling because after selecting a "chunk" the investigator goes to great effort to employ good probability sampling procedures within the "chunk."

To illustrate, let us suppose that a state health department wishes to determine the magnitude and nature of the problem of home accidental injuries in the state. A certain county is selected because it is believed by those who know the state well to be typical of the living conditions in the state—about average in urbanization, level of living, etc. An additional reason for its selection is that the county has an energetic county health department, quite well supplied with nurses who can be called upon to do some field work. Within the county a probability sample of areas is drawn, households are listed for the sample areas, and a subsample of households is chosen at random. By means of interviews with housewives in the sample households the desired information is secured and analyzed. Population rates are computed specific for age and sex and these are multiplied by the latest estimates of the state population for the corresponding groups. All goes well until the question arises: What sort of confidence interval can be used in connection with this state estimate? How far off might it be because we have interviewed in only 1,500 households instead of every household in the state?

At this point there is nothing to be said except that we do not know. Was it worth while to go to the trouble of probability sampling within the county? We can make statements about the sampling error of an estimate for the county, but the real interest lay in obtaining some knowledge of the situation in the state as a whole. A probability sample of 1,500 households drawn from the population of households in the state is very fine to talk about, but in this case, perhaps, confining the study to the county "chunk" meant the difference between a project costing $2,000 and one costing $20,000. (The tenfold difference is not unreasonable if we suppose that a great deal of help could be obtained in the county without extra cost to the state.) This type of situation arises very frequently in public health work of all kinds. What is the answer? Do we take a "chunk" and trust to our imperfect knowledge of the variability between "chunks" of this sort to decide how much faith to put in the estimates? Or do we spend the money for a sample of the entire population in which we are interested so that sampling error of estimates will be "controllable"?

The kind of guidance that the statistician can give in attempting to answer this question is far from perfect. There are some deceptive answers that the statistician can warn against, though possibly the warning is not needed. First, it is unrealistic, having selected a convenient "chunk" for study, to look for the population of which this "chunk" can be considered a good sample. Second, it does not help much to decide that study of the "chunk" is to be considered purely preliminary—a pilot study, unless there is a reasonable chance that the pilot study will be followed up by a more conclusive study later. The statistician should be the last to minimize the importance of pilot studies. Nevertheless, there comes a time when we have to decide whether to proceed to a full-scale sampling (or complete coverage) investigation or drop the project for lack of funds. Finally, demonstration that the "chunk" is similar in many respects to the population about which we should like to generalize may increase our confidence in the estimate, but it still does not make it possible to attach a confidence interval to the estimate. There are instances on record when there appeared to be excellent agreement in a number of characteristics between a nonprobability sample and the population from which it was drawn; yet the confidence to which this agreement gave rise was later found to have been misplaced because the sample and population differed in one respect, namely, the characteristic the sample was designed to measure!

About the most that the statistician can say regarding the advantage of probability over "chunk" sampling is this: The use of probability sampling will often raise the cost of an investigation as compared with a study based on those parts of the population that can be reached most conveniently, or to which there is easy access because of the cooperativeness of some official, or which combine convenience with an appearance of representativeness. However, the additional money is being spent upon a very valuable commodity—a quantitative statement concerning the limits of error resulting from sampling. The marketability of the final product, the likelihood that the results will be trusted and used, are enhanced greatly by the addition of that statement. Sometimes the extra cost of doing a probability sampling job is trivial; sometimes it is completely out of reach. Whether the additional quality of the product is worth the additional cost must finally be decided by the purchaser. The statistician's job is to be a wise agent, presenting as accurately as possible the comparative costs and as much data as are available about the variability of the quantities to be measured.

CONCLUSION

It has not been possible in this report to do more than touch upon the principles of sampling, the determination of the sample size necessary to insure a specified degree of sampling accuracy, the variety of sampling designs that are possible, etc. For a thorough treatment of these subjects the reader is advised to consult the list of references. A short time ago an apology would have had to be made for the fact that discussion of the application of sampling to surveys and experiments could not be found in books. The health department statistician who wished to increase his knowledge of this branch of statistics in preparation for a particular job would have had some difficulty in bringing together the papers bearing on this subject. Much of the practical knowledge had never been put down in writing. Recently, however, some excellent books have appeared. The names of several of these (see references 2, 4, 7, and 8) and other papers and pamphlets that may be useful to the public health statistician have been included in the references.

REFERENCES

1. United Nations Sub-commission on Statistical Sampling. A Brief Statement on the Uses of Sampling in Censuses of Population, Agriculture, Public Health, and Commerce. (Prepared by W. Edwards Deming, member of the Sub-commission.) Lake Success, N. Y.: United Nations Publications Sales Number: 1948 XVII.1 (Feb.), 1948.
2. Deming, W. Edwards. Some Theory of Sampling. New York: Wiley, 1950.
3. Cochran, William G.; Cornfield, Jerome; Hansen, Morris H.; and Hurwitz, William N. Modern Methods in the Sampling of Human Populations. A.J.P.H. 41, 6:647–668 (June), 1951.
4. Hansen, Morris H.; Hurwitz, William N.; and Madow, William G. Sample Survey Methods and Theory. Vol. I. Methods and Applications. New York: Wiley, 1953.
5. Field Methods in Sample Interview Surveys. Report of a Conference Held at Ann Arbor, Michigan, March 22–24, 1951; Sponsored by the Survey Research Center, University of Michigan. September, 1951.
6. Berkson, Joseph; Magath, Thomas B.; and Hurn, Margaret. The Error of Estimate of the Blood Cell Count as Made with the Hemocytometer. Am. J. Physiol. 128, 2:309–323 (Jan.), 1940.
7. Cochran, W. G. Sampling Techniques. New York: Wiley, 1953.
8. Yates, Frank. Sampling Methods for Censuses and Surveys. New York: Hafner, 1949.
9. Patton, Robert E. The Sampling of Records. Pub. Health Rep. 67, 10:1013–1019 (Oct.), 1952.
10. Commission on Chronic Illness. Guide to Making a Survey of Persons in Establishments Which Provide Nursing and Personal Care. Processed. Oct., 1953. See special supplementary statement entitled "Sampling."

BIBLIOGRAPHY

History of Sampling

Stephan, Frederick F. History of the Uses of Modern Sampling Procedures. J. Am. Statist. A. 43, 241:12–39 (Mar.), 1948.

Two Pamphlets on Sampling

McCarthy, Philip J. Sampling. Elementary Principles. Bull. No. 15. New York State School of Industrial and Labor Relations. Ithaca, N. Y.: Cornell University (Apr.), 1951.

Hendricks, Walter A. Mathematics of Sampling. Special Tech. Bull. Virginia Agricultural Experiment Station, Blacksburg, Va. Feb., 1948.

Examples of the Use of Probability Sampling in Public Health

Hoyt, Cyril J.; Knutson, Andie L.; and Derryberry, Mayhew. What the People Know. Monthly Bull., Indiana State Board of Health, November and December, 1950.

Two random samples of the adult population of the city of Mishawaka, Ind., were interviewed, one before and the other subsequent to a tuberculosis x-ray survey in order to evaluate the effect of the survey and of an information campaign upon public information and attitudes toward tuberculosis.

Horvitz, Daniel G. Sampling and Field Procedures of the Pittsburgh Morbidity Survey. Pub. Health Rep. 67, 10:1003–1012 (Oct.), 1952.

An area probability sample of households located in a part of the Arsenal health district in Pittsburgh was used in conducting a survey which had as its objectives: (1) to provide general measures of the health status of the household population through the collection of data on recent illness, hospitalization, and accident experience; (2) to characterize the population under study through the collection of basic social and economic information; (3) to provide a general picture of the pattern of demand for health services; and (4) to furnish the nucleus of families to be followed for long-term studies.

Hemphill, F. M. A Sample Survey of Home Injuries. Pub. Health Rep. 67, 10:1026–1034 (Oct.), 1952.

The primary purposes of this study were to find the rate of accidental injuries occurring in the homes of Washtenaw County, Mich., and to study selected factors—environmental, physical, and psychological—associated with injuries and injured persons. Stratified area sampling was used.

Abdellah, Faye G., and Levine, Eugene. Work-Sampling Applied to the Study of Nursing Personnel. Nursing Res. (June), 1954.

Work sampling in industrial management dates back to the 1880's. This paper tells of the application of work sampling technics to estimating the distribution by type of activity of the working time of nurses in hospitals. The method has general applicability to time studies of other categories of personnel.

Report on Market-Milk Supplies of Certain Urban Communities. Pub. Health Rep. 53,

32:1381–1399 (Aug. 12), 1938. See particularly Appendix entitled "Sanitation Ratings of Milk Sheds," by Frank, L. C.; Fuchs, Abraham W.; and Dashiell, Walter N.

This describes a widely used schedule of sampling which has as its objective the rating of communities on the extent to which they have enforced sanitation requirements for pasteurized and raw milk. The sampling units are raw milk dairy farms, farms delivering milk to pasteurization plants, and pasteurization plants. A similar sampling plan has been used for sampling restaurants.

Hoffer, Charles R., et al. Health Needs and Health Care in Michigan. East Lansing, Mich.: Michigan State College, Agricultural Experiment Station, Section of Sociology and Anthropology. Special Bull. 365 (June), 1950.

This describes a stratified random sample containing a representative cross-section of all households in Michigan, exclusive of Wayne County, and a random subsample of all adults in these households. The purpose was to measure the extent of unmet need for medical attention, the opinions and attitudes people have regarding health facilities, the extent of participation in prepayment plans for medical and hospital care, and the people's awareness of local community health needs.

Taback, Matthew, and Williams, Huntington. Statistics in a Health Department Medical Care Plan. Pub. Health Rep. 68, 2:157–166 (Feb.), 1953.

Sampling is used in this program to reduce the cost of summarizing some of the medical records from the plan, for example, the clinic histories of adults presenting themselves for general examinations, and drug invoices. Results are used to improve administration of the plan.

Monthly Vital Statistics Report. Annual Summary for 1952. Part 2—Estimated Numbers of Deaths and Death Rates for Selected Causes: United States, 1952. (Based on the Returns of a 10-Percent Sample of Death Certificates Received in State Vital Statistics Offices in 1952.) Vol. 1, No. 13. Washington, D. C.: National Office of Vital Statistics, Public Health Service, Department of Health, Education and Welfare. July 21, 1953. See also Instructions to Special Agents for the Current Mortality Sample. National Office of Vital Statistics, January, 1952.

These publications describe a sampling project that is designed to provide preliminary national statistics on mortality by cause-of-death, age, and sex. The data are available monthly within about two months of the end of the month to which they apply.

Committee on the Hygiene of Housing, American Public Health Association. An Appraisal Method for Measuring the Quality of Housing: A Yardstick for Health Officers, Housing Officials and Planners. Part I. Nature and Uses of the Method. New York: American Public Health Association, 1945.

Although the purpose of this publication is mainly to describe the objective rating system and its application to the planning of housing programs, Chapter II does describe briefly the use of a stratified random sample of dwelling units in New Haven, Conn., by the New Haven Health Department and the City Plan Commission. When some action has to be taken on the basis of the condition of a single dwelling unit, the sampling method is, of course, not applicable. On the other hand, when results are to be used to plan or evaluate a housing program, the sampling method used here or one similar to it should achieve great economies.