IMPROVEMENT OF METHODS USED IN THE HIGHWAY PLANNING SURVEYS

By C W Vickery
Associate Statistician, Public Roads Administration

SYNOPSIS

The emphasis in this report is on observational methods rather than on methods of statistical analysis. The main problem considered is that of obtaining a maximum amount of information for a given sum of money. A unit introduced by R. A. Fisher is used to measure the quantity of information per dollar provided by an observational technique used, or proposed for use, in the highway planning surveys. The chief function of the statistician should be to design unbiased and efficient methods of observation; calculation of statistics is of secondary importance.

Improvement of observational technique has resulted in (1) the development of special machines for recording traffic counts and other data, particularly data on traffic behavior, and (2) the application of scientific sampling methods to traffic counting and interview surveys. Interviews from only one per cent of the vehicle owners have provided surprisingly reliable information about road use. Economy has led to the use of traffic counts as short as five minutes on city streets. A newly designed method of moving traffic enumeration reduces the length of the count at each point to zero, it may be useful in armed reconnaissance of hostile territory from airplanes

An improved type of origin-destination survey has been designed for detecting traffic potentials that demand new highway facilities and predicting the volume of traffic that would use a proposed new facility If conducted on a State-wide, or Nation-wide, basis, this study may help to place priority of new projects above individual judgment and political pressure.

The emphasis in this report is on observational methods rather than on methods of statistical analysis. The problem in practical research is that of securing a maximum amount of information for a given sum of money. It is this consideration that determines the design of observational technique, which in turn determines the statistical methods that are appropriate. The cost, and sometimes the possibility, of collecting data and the knowledge available for control purposes determine the observational designs that are available; the theory of the design of experiments selects from the available designs the one that provides the maximum amount of information for the money. Quantity of information does not depend solely upon the number of observations but also upon the manner of making the observations and the methods of statistical analysis. Statistics can never improve upon the observations after they are made; biased statistics may distort, and inefficient statistics may waste, the information supplied by the best arranged experiments.

In certain branches of practical research, however, it is usually impossible to control the conditions of the observations; all that can be done is to observe phenomena under the conditions that nature presents. This circumstance has led to the attempt to substitute statistical control for experimental control. The use of such methods sometimes makes it possible to estimate, and allow for, the effect of a variable quantity which it is impossible to keep constant. This accounts for the complicated statistical methods employed in astronomical, biological, and economic research.

Natural knowledge, in whatever field, increases through the formation of hypotheses by the intuition, the elaboration of them by logical processes, and the testing of them through observation. While pure logic and mathematics are concerned with both intuition and reasoning, the science of statistics has as its peculiar domain the combination of the results of observation and the discrimination of their significance. Statistics must, therefore, distinguish unbiased from biased methods of observation and appraise their efficiency. It is required of a statistic that it be unbiased and it is desirable that it be efficient; these qualities in a statistic require that the observations on which it is based be made in an unbiased and efficient manner.

Many of the statements made in this paper require complicated mathematics for their proof. References to literature have been given wherever possible; in a few instances statements depend from unpublished work of the author. Frequently the author has given his opinion on questions without any pretense or hope of being able to give a rigorous proof of the statements. This state of affairs is uncomfortable but it appears inevitable in the present state of development of highway planning survey methodology.

Publications and reports referred to have been listed at the end of this paper and references are indicated in the text by numbers enclosed in parentheses.

Traffic Volume Studies

METHODS BASED ON MANUAL COUNTING

The technique of making surveys of street and highway traffic has shown an interesting development within a short period This development has included methods of collecting data, sampling methods, and the statistics appropriate to these sampling methods. Until a few years ago all traffic counts had to be made, through manual tallying or through the use of some form of manually operated counting machine. While in a few instances this work of counting was done incidentally to the collecting of tolls at bridges or tunnels, in most cases the taking of manual counts was an expensive operation The necessity for economy was the driving force in the production of sampling methods and statistics which made it possible to do with much less than continuous counting. This development has taken two distinguishable forms which we may call the traffic pattern method and the quadrature method. The necessity of economy has also led to the development of machines for counting traffic and recording the counts automatically. Improvements in methods of collecting data are leading to modification of the sampling and statistical methods.

Traffic Pattern Method

The goal of the traffic pattern method is to make a single short count suffice for an estimate of the traffic on a facility for the period of a year The short count may cover a week, a day, eight hours, a single hour, or even ten minutes, but the point is that a single continuous count is taken and from this an estimate of the annual traffic is made The practical advantages of such a method, if successful, are obvious.

In order to carry out such a process of estimation, or even to decide whether we can use such a process, it is necessary to investigate the patterns of traffic variation over the day, the week, and the year. Fundamental work in the study of traffic patterns has been done by Nathan Cherniack (3), results of which were presented to the Highway Research Board in 1936. Based on such studies is what is here called the traffic pattern method In accordance with this method we prepare an estimate of the annual traffic by multiplying a single count by an expansion factor. The expansion factor is obtained by finding the ratio of the annual traffic to the traffic during the period of the short count. This factor is usually calculated from data obtained at another facility where a continuous count has been taken for a year, or a series of short counts spaced over the year.

There are two types of errors that may be involved in such estimates. Errors in the short count, and errors in the expansion factor. Errors of the first kind have been commonly recognized, but errors of the second kind have frequently been overlooked. Errors of the second kind appear when we "borrow" a traffic pattern (to use Cherniack's term) from another facility. The borrowed pattern may be entirely wrong and it will almost never be entirely right. The worst part of the situation is that we usually have no way of estimating the probable error introduced by the expansion factor. It is necessary to investigate both sources of error before we can produce by this method an estimate with any claim to being called scientific.

Quadrature Method

Another way of estimating traffic volume is the quadrature method This method substitutes a series of short counts scattered over a year, or other period, for a continuous count over that period It is called the quadrature method because it is analogous to the method of quadrature commonly employed by engineers to estimate the volume of an excavation Fundamental work on the reliability of estimates prepared from a series of counts, of various durations and with different systems of spacing, has been done by W. Arthur Shelton (29).

The work of Cherniack (3), and others, particularly the work done in the highway planning surveys, has demonstrated that traffic volume is subject to systematic diurnal and seasonal variations. The result is that in sampling such events we have errors of quadrature as well as random errors to contend with Periodic spacing of the counts will reduce to a minimum the error of quadrature and also insure a random sample from the residual variations

At all events, periodic spacing of the counts is the correct procedure if we know nothing about the periodic variations of the traffic stream or the dispersion of the residual variations If we do have such knowledge we may be able to improve our sampling method. Cherniack has noted that the time of the year of average traffic is also the time of minimum dispersion If we know for a particular road section when the month or week of average traffic is, that would appear to be the ideal time to do all the sampling, provided average traffic is all that we are interested in. If we are also interested in peak season traffic, we will have to take some counts during the peak season, and we may find that the counts need to be spaced more closely than in the average season because the dispersion is greater Cherniack, at any rate, has noted higher dispersion both in the high and in the low traffic seasons than in the average traffic seasons

There is some further qualification to the statement that the counts should be periodically spaced It is now known that Saturday traffic, as well as Sunday traffic, often exhibits greater dispersion than weekday traffic. Hence we need to sample differently the three types of days, Saturdays, Sundays, and weekdays. The indications are that weekdays, except for holidays, may be treated alike in most instances. The allocation of the number of counts to Saturdays, Sundays and weekdays is to be determined by the principle of stratified sampling Once we have determined the number of Sunday counts to be taken, these counts are then spaced periodically over the year. A similar statement applies to the weekday and Saturday counts A further refinement would be possible in cases where the information is available, but usually would be out of the question; if we know that the dispersion varies seasonally, than we should space the counts closer together during the season of greater dispersion This is a refinement which is not usually to be recommended.

Method of Extremely Short Counts

The author (42, 43) has extended the quadrature method to the estimation of the traffic during a particular day by replacing a continuous count for the day by a series of short counts spaced periodically over the day For estimating the annual traffic it is assumed that a series of days spaced over the year would be used. The method of extremely short counts in combination with the pattern method has been used by R. O. Swain (32, 33) for estimating city street traffic

The lengths of the short counts and the frequency of their spacing depend upon the volume and variability of the traffic The more nearly the traffic approaches a continuous flow, the shorter can be the individual counts R O. Swain, assisted by the author, used 5-min counts spaced 30 min apart at Amarillo, Texas The lengths of the counts and their spacing were determined largely by practical judgment based on observation of local conditions In places where a study has already been made of traffic patterns, it would be possible to time the counts on a more nearly scientific basis.

The question has naturally arisen as to whether this method can be applied to surveys of rural highway traffic The author feels that there are practical difficulties in the way of such an application except under suburban conditions resembling urban conditions The main difficulty is that of transporting the traffic observer from one station to another between counts. In urban conditions, with stations only a few blocks apart, the lost time and transportation cost are not serious obstacles At Amarillo lost time and transportation cost were largely eliminated by Swain's method of placing traffic observers on the tops of the taller buildings from which they could see a number of intersections with little or no change in position. On a rural highway system, however, with stations spaced ten miles or more apart, the problem of transportation time and cost in getting the observer from station to station becomes an important obstacle. A series of 1-hr counts, or 15-min counts, could be spaced at regular intervals over the year, these counts being suitably distributed for day of the week and hour of the day Thus ample time could be allowed for getting from station to station, and if enough of the counts were secured, estimates of the desired degree of reliability could be provided The practicality of the method appears to be limited to urban and suburban conditions

METHODS INVOLVING AUTOMATIC TRAFFIC RECORDERS

The development of automatic counting and recording devices has made it possible to obtain continuous traffic counts at a much lower cost than was possible by manual methods. The development of these machines brings with it changes in the sampling and statistical methods.

Since these machines are particularly efficient for securing continuous counts of 24-hr. or longer, it appears that with them the 24-hr. day can be taken as the sampling unit. This eliminates all worry about diurnal variations. The problem of estimating the annual traffic may thus be stated as the problem of spacing a series of weekdays, a series of Saturdays, and a series of Sundays over the year. It must not be supposed, however, that the value of a continuous count is proportional to its length, for purposes of statistical estimation

Use of Fixed Type Recorders

The fixed type traffic recorders, besides furnishing highly reliable traffic volume figures, also furnish laboratory data for statistical investigations. Analysis of variance studies, made by the methods of R. A. Fisher (11), are capable of indicating the proper distribution of counts among weekdays, Saturdays, and Sundays and which seasons of the year need the heaviest sampling

Patterns of diurnal, weekly, annual, and long-term variations have been studied considerably on the basis of data from fixed type recorders and need further investigation These data also make possible the study of peak season traffic, peak hour traffic, and periods of congestion generally

DISTRIBUTION OF TRAFFIC OBSERVATION STATIONS

Whether counts are secured manually or by automatic recorders, the question arises of how to locate stations along a road for counting traffic. Of course, the correct way of locating the stations depends on our purpose. We may be interested in knowing the traffic volume only at the intersections or only at railroad grade crossings, at bridges, or at unincorporated towns; in that case, these points, and no others, are the ones at which the counts are to be taken. In the highway planning surveys, however, we are interested in knowing the approximate number of vehicles that go past each point on the road during a given period, such as a year Our problem is that of so spacing the stations along the road as to produce this result with the greatest degree of reliability

What is frequently overlooked is that in studies of this type there are errors of quadrature in addition to random errors. By a random location of stations we may greatly increase the error of quadrature without diminishing the random errors. In fact, periodic sampling is frequently the best way to insure a random sample from the random errors; it is certainly the best way to reduce errors of quadrature. This is not to deny that the best way to sample from an ordinary statistical distribution is to select the elements of the sample at random What is contended is that this method should not, because of its applicability to one type of data, be applied to types of data to which it is not applicable, or for which there may exist a better method.

Counting traffic at intersections has the advantage that intersections are likely to be points of sharp change in the traffic profile. Furthermore, if manual methods are used, counts can be secured on more than one roadway without additional cost For certain purposes of traffic engineering the volume of traffic in the various turning movements at intersections is of more importance than the traffic volume along the road If it is decided that counts should be taken at or near intersections, this does not vitiate the principle of periodic spacing of the stations If there are numerous intersections to choose from, the proper periodic distance can be measured on the odometer and the next intersection designated as a count station. If there are other intersections which, because of their importance, need to be counted, this will not cause any difficulty provided the counts are properly weighted by length of road section in calculating vehicle miles Suitable application of the principle of periodicity might be made by taking a count at every second intersection, or every tenth intersection If intersections are far apart, stations may be spaced periodically between intersections

Traffic Behavior Studies

METHODS OF COLLECTING INFORMATION

Methods of studying traffic behavior have shown an interesting development within a short period. Five years ago, or even less, the study of traffic behavior was in what we might call the speculative phase From certain assumptions about the speed and spacing of vehicles it was found possible to develop formulas for highway capacity, the vehicle-minutes of delay produced by stoppage of traffic, and other important characteristics of motor-vehicle traffic Such a procedure is just what one would expect in accordance with scientific methodology; the only difficulty is that many of the assumptions were not correct and thus the conclusions were inapplicable. The fault was not that these early investigators theorized, but that their theories were not based on statistics calculated from observational data

The study of traffic behavior suffered only a few years ago from paucity of observational data and from over-simplification of the theories. It was necessary to find a way to collect sufficiently complete information about the various aspects of traffic behavior relevant to road design and still stop short of a motion picture of the traffic on the road. This is what special equipment designed in the Public Roads Administration undertakes to do This equipment, as well as its utility, has been described in papers published by E. H. Holmes (14) and E. H. Holmes and S. E Reymer (15). Results of studies using this equipment have been presented to the Highway Research Board and published by O. K. Normann (26)[1]. The result is that today these studies give promise of becoming one of the most notable examples of the application of scientific methodology to highway planning survey work.

Certain of the earlier devices for collecting information about traffic behavior were unsatisfactory because they introduced bias into the sample collected Studies were made, for example, of passing practices, by an observer in a motor car traveling in the stream of traffic he is undertaking to study. Such a study would necessarily be biased because of the influence of the observation car. The Eno speed detector has to be set up beside the road in a rather conspicuous position and is liable to cause reduction in speed out of curiosity or out of fear of law enforcement. The usual necessity for measuring the speed of only a sample of the traffic may lead to bias unless care is exercised; there might easily be a tendency to measure too great a proportion of the slow vehicles because of their longer exposure to view Care has been taken in the design of the recent equipment to avoid these biasing tendencies

With the accumulation of masses of highly detailed information about traffic behavior, it is to be hoped that theorizing will not be abandoned. The older theories proved inadequate because many of their assumptions were not realized in practice. It would be easy to go to the opposite extreme of collecting data, classifying them, and calculating statistics without anything further Practical utilization of information requires the formation of theories based on statistics calculated from the information.

PROBLEM OF CONGESTION

The earlier speculative studies of traffic behavior usually assumed that all vehicles were traveling at the average speed or at a "normal" speed This assumption of uniformity of speed for all vehicles is not satisfied in practice and it leads to estimates of highway capacity that are too high for practical conditions. In fact, one of the most significant characteristics of motor-vehicle traffic is that the vehicles do not all travel at the same speed. The tendency of vehicles to travel at the same speed has been found to be a good index of congestion Complete congestion results when all vehicles are required to travel at the same speed

One simple study of highway congestion that can be made is to classify the hours by traffic volume groups. We can then determine the number of hours, and also the number of vehicles, that are exposed to a given or greater traffic volume We are expected to provide facilities adequate to all but a small percentage of the higher volume hours. Graphs prepared to show the cumulative distribution of hours,

[1] Also see this volume, pages 379-392.

or vehicles, by hourly volume groups, can be adapted to different total volume groups in the vicinity of the point where the counts were made by a device suggested by the author The graphs are drawn with a logarithmic scale for the hourly volume groups and a slide is provided with the same scale By sliding the movable scale forward or backward along the first scale it is possible to read the graph for lower or higher total traffic volumes Such a device should not be used unless we are confident the distribution is the same for the different total traffic volumes

Origin-Destination Studies

In the Highway Planning Survey work a distinction has been made between special origin-destination studies and general origin-destination studies. Special origin-destination studies are designed for throwing light on a special problem in a limited locality The problem may be that of providing a by-pass route around a town, providing more adequate routes for traffic through a city, or building a more direct route between two points on the rural highway system A frequently used technique for making these studies is to stop the traffic for interviewing at a number of strategically located points near the situs of the problem. The technique for making these studies and analyzing the data has been rather well developed through repeated performance

NEED FOR A NEW TYPE OF GENERAL ORIGIN-DESTINATION STUDY

In a comprehensive program of state-wide highway planning, it appears that there would be an important place for general origin-destination studies which would collect something like the detailed type of information collected in the special origin-destination studies Frequently it is easy to build up a case for the construction of one particular facility that has been proposed A special origin-destination study is made and benefits to the traveling public are estimated which appear ample to justify the building of the new facility However, the question arises as to whether there may not be other equally meritorious facilities whose construction has not been proposed or not advocated by a sufficiently influential spokesman to bring them to the point of serious consideration. There may be such facilities which have not been adequately considered but which would be more amply justified on the basis of social benefits than the one under consideration This is a state of affairs which the highway planning surveys have undertaken to eliminate. By collecting information on a state-wide and nation-wide basis for all types of rural highway facilities the goal is to establish priority for new highway improvements whose basis will be economic in character and independent of any sort of agitation by special interests It appears that the attainment of this goal would require state-wide and nation-wide origin-destination studies of a type designed to provide information with regard to the traffic pressure, or traffic potential, between any two points on the highway system.

Information Needed for Proposed New Study

In the proposed type of general origin-destination study the information of first importance is the route traversed by each vehicle. The determination of the route makes it possible to determine the proximate origin and destination of each vehicle with respect to every point that the vehicle affects There appear to be at least two very different ways of collecting the data for a study of this type One of these is what we might call the origin-destination method, and the other we might call the road use method The origin-destination method would simply extend the special origin-destination technique to a State-wide basis. Traffic would be stopped at stations located on the highway system and brief interviews would be secured pertaining mainly to the highway route traversed by the vehicle in reaching and leaving the station It is contemplated that the route of the trip would be traced to its ultimate origin and destination or to its point of entering or leaving the State These stations could be operated at periodic intervals much as the loadometer stations were operated

In the road use study the drivers are expected not only to record the origins and destinations of their trips, but also to indicate by number each highway used as well as the point of entering and leaving the highway From this information it is possible to trace on a map that portion of each trip that was made on the State highway system. The road use study thus provides general origin-destination information of a highly detailed kind whose value for this purpose has not yet been utilized except in a special study made by the Public Roads Administration in cooperation with the University of Wisconsin In considering the value of road use information for this purpose, the author is skeptical about the reliability of the individual questionnaires It appears almost too much to expect that an ordinary individual can describe with much accuracy the various trips that he has made over a year This statement applies to a less extent to trips made on the State highway system, particularly the longer trips

Analysis of Data

Whether this type of information is collected by the origin-destination method or by the road use method, a somewhat complicated punched card technique for the analysis of this information has been proposed by the author (36) This method involves punching, for each trip, on a tabulating card certain control information, an expansion factor, and the number of each highway traversed together with the station of entry and station of departure on that highway By proper sorting and tabulating of the cards, involving totaling the expansion factors, estimates can be prepared of the traffic density at each point on the highway system, as well as the amount of through traffic between any two points A simple modification of the technique will provide estimates of the through traffic over any given section of highway This method can be extended to prepare estimates of the through traffic between any two points on the highway system regardless of whether they are situated on the same highway or not. It appears that this information would be of value in estimating the traffic potential between any two points on the highway system and thus attempting to estimate the amount of traffic that one would expect to be diverted to a new facility joining the two points.

A difficulty presents itself in the analysis of information collected by the origin-destination method which does not result when the information is collected by the road use method A similar difficulty was encountered in using origin-destination information to estimate the traffic at points on the county road system where counts were not taken The method of extensional analysis (35) was designed to take care of this difficulty which is that if we pick up trips at a point on a road, the longer trips are more likely to be picked up, and the factor of bias is proportional to the length of the trip The result is that if we are trying to estimate the average trip length, we should use the harmonic mean rather than the arithmetic mean. Furthermore, in attempting to estimate traffic density and traffic potential we need to make a similar correction for bias

Road Use Studies

Road use studies have considerable similarity to the public opinion polls that have become increasingly popular recently The road use studies, however, present more complicated statistical problems because their questions require quantitative answers rather than a simple yes or no The result is that the road use studies require what Arne Fisher (8) calls heterograde statistics as distinguished from homograde statistics, which are required in the public opinion polls The greater complexity of the road use studies necessitates a more complicated design of the sampling method, as well as more complicated statistics.

METHODS OF COLLECTING INFORMATION

The road use studies furnish an excellent example of the principle that a usable sampling method is dependent upon the methods of collecting data that are available or economical It is easy enough, provided one is acquainted with contemporary statistical literature, to design a sampling method for road use studies which would be entirely scientific. The difficulty, comes in collecting the information. Three different methods of collecting road use information, which have either been used or proposed for use in planning survey work, will be discussed. These may be called (1) the personal interview method, (2) the school method, and (3) the reporter method The method of questionnaires by mail has not been included, because the unreliability of this method has been abundantly demonstrated.

Personal Interview Method

The personal interview method has been used in a number of States and apparently it was the method originally contemplated when the road use studies were designed It consists of having an interviewer call on the motor-vehicle owners, either at their homes or places of business, and fill out the questionnaire forms through asking questions and recording the answers.

The first question which presents itself is, how are we going to determine the individuals from whom we shall attempt to secure interviews. The most nearly scientific method available would be to select a random sample of owners and their addresses from the motor-vehicle license numbers, making use of a table of random sampling numbers such as that of Kendall and Babington Smith (18) or the older table of Tippett (34). The problem then is how to secure a questionnaire from each of these individuals. An almost certain result is that the interviewers would not be able to locate all the individuals on the list. Of course, we can always secure interviews from other individuals to take the place of individuals on our list from whom we are unable to secure interviews. However, precisely the same result would have been accomplished by making up a longer list to start with. It is conceivable that people who use their motor cars a great deal each year will be more difficult to make contact with because the probability is greater that they are in their motor cars when the interviewer calls

The danger of introducing bias lies not in the possibility that we may not secure enough interviews, but in the possibility that we may not be able to secure interviews from everyone on the list, whatever its size, and that the people omitted will fail to be a random sample and thus have a biasing influence There is apparently no satisfactory method for getting around these difficulties. The best that can be done is to have the interviewers thoroughly trained, for them to exercise persistence in follow-up work, and for us to be prepared to spend the money it takes to get complete, or nearly complete, coverage.

Frequently, however, the interviewer does not start out with a list of names and addresses In the road use work in some of the States it is said that the interviewer merely drives along the road, stopping at houses as he comes to them. This method might be entirely satisfactory provided it were properly designed and carefully executed. The result would be unfortunate if the interviewers were under pressure to secure as many interviews as possible while traveling as few miles and spending as little money as possible Even if they were required to secure a certain percentage of the interviews on the county road system, it would be inevitable that these interviews would be secured from residents close to town and close to the primary highways.

The Reporter Method

In an effort to avoid some of the difficulties inherent in the personal interview procedure, a method has recently been proposed which is here called the reporter method This has not yet been used in highway planning survey work, but similar methods have been used by the U. S. Department of Agriculture and by the U S. Weather Bureau. The method consists basically of selecting a small group of citizens to serve as reporters. It is assumed that the list of reporters can be selected in accordance with scientific principles and can be fairly small, perhaps as small as one percent or 1/10 of one percent of the motor-vehicle owners, perhaps even smaller. The list of reporters might be compiled by suitable drawing from the motor-vehicle registration records or by periodic sampling carried out by an interviewer driving about the country The main point is that the number of reporters would be small and they would be selected with the utmost care. The reliability of the results would be improved by keeping the same group of reporters year after year Of course, this would not be strictly possible because of death, migration, and other changes that befall human beings.

Not only does this method attempt to secure a well-distributed sample, but also it undertakes to secure more accurate questionnaires. It appears that the weakness of the road use surveys, as previously conducted, lies not so much in improper distribution of the sample as in unreliability of the individual questionnaires. The reporter method tends to get around this difficulty by inducing the reporters to keep records of their driving as they go along, instead of waiting until the end of the year and then trying to recall it.

It appears that the difficulties involved in this method consist in finding a group of reporters and stimulating their interest sufficiently to produce the reports The author would be skeptical of the possibility of this if it were not for the success that has been achieved by the Department of Agriculture and the Weather Bureau

The School Method

The school method of collecting road use interviews has been used in Texas and other States. This method has the advantage of bringing in a great number of questionnaires at low cost per questionnaire. It has also been claimed that it produces more reliable questionnaires. Of course, this method makes no pretense of securing a random sample from the motor-vehicle owners or, in any sense, of being a scientifically drawn sample. The interviews secured come almost entirely from that class of motor-vehicle owners who have children of high school age It is conceivable that this subclass of motor-vehicle owners is representative of the entire population of motor-vehicle owners, but it has never been demonstrated.

In studies of this type too much emphasis has been placed on the size of the sample and too little on its proper distribution. Having secured the sample, one feels that it must be used, but the cost of analysis may turn out to be prohibitive. When a large sample has been obtained it may turn out to be better to select a small subsample for careful analysis. That is what was done in the Texas road use survey.

It may be possible to draw a small subsample which is better distributed with respect to the population than the original Making use of the method of stratified sampling, we can calculate the number of questionnaires that should be secured from each stratum of the population By sorting the original large sample in accordance with this method of stratification, we can draw from each stratum of the large sample the proper number of questionnaires and throw the rest back. Now we have no right to assume that the original large sample produced by the school method is a random sample from each stratum of the population, and hence no matter how careful we are to draw a subsample at random with respect to the original sample, we have no justification for assuming that the subsample will be random within each stratum of the population. In this respect we are unable to improve upon the original sample, and the best we can do is to avoid making matters worse by introducing new components of bias. Once the small subsample has been drawn, special statistics are calculated from it as provided by the method of stratified sampling

SAMPLING METHODS

Whatever method may be used to secure the interviews, there appears to be no question that stratified sampling is the most efficient method to use. The only question is how best to apply this method In practice we are never able to carry out the theoretically most desirable system of stratification. We are always limited by the information that is available for control purposes In the cases where desirable control information is not already available an attempt is sometimes made to secure such control information by a process of sampling This sampling procedure is independent of the study that we really want to make, and it is carried out in advance of starting that study. The problem thus presented is that of double sampling, and the question of the proper distribution of available money between the large control sample and the final small sample has recently been solved by J Neyman (24). Such a situation was presented by the motor-vehicle allocation study and the road use study, but nowhere were the advantages presented by the method of double sampling fully utilized It is now possible to state what the allocation of funds should be between M.V A and road use and what size of samples should be obtained for each study to insure a required degree of reliability.

RELIABILITY OF RESULTS

In all the work that has been done in the motor-vehicle allocation and road use studies, almost no attention has been given to the reliability of the results. Granting that the sample has been properly designed and the design carefully executed, then it is possible to calculate measures of the reliability of the various means and totals that have been presented in tabular form. The result is that we are never able to say that the correct answer is a definite number. The most we can say is that the correct answer probably lies between two limits which we state, and we can state what this probability is.

There was, indeed, a limited study of the reliability of road use statistics made in Texas by representatives of the Public Roads Administration working in cooperation with the State personnel (11a). This study is an example of the sort of thing that should have been done in all the States

Road Life Studies

ACTUARIAL PRINCIPLES, WINFREY'S WORK

The basic principles of actuarial science applicable to human life contingencies have been developed to a high degree since the time of Halley, who not only discovered a comet but also constructed a mortality table. Attempts have been made to apply statistical methods to all manner of business problems, including the prediction of futures, but the important point is that actuarial science applied to life insurance has been eminently successful, in contrast with numerous other attempted applications It is relatively simple to make up a mortality table from vital statistics such as those collected by insurance companies and departments of public health It takes much care to have the table reflect the actual state of affairs, but such tables have been being made for at least two centuries and the technique has been rather well developed

When such a table has been prepared, however, it represents merely the distribution of mortality as observed in a sample of human lives or of industrial property retirements If we prepare a mortality curve, or a survivor curve, we find that it is irregular, whereas it is expected that such a curve for an infinite, or a large finite, population would be relatively smooth The operation of smoothing a jagged curve in order to obtain a curve which, it is hoped, will more nearly represent the state of affairs in the population, is called graduation. Perhaps the most nearly scientific method of graduation employed at the present time is the process of curve fitting There are, however, two problems involved in curve fitting which require exercise of judgment, which vitiates the scientific character of this procedure. In the first place, it is necessary to decide what type of curve to fit, and in the second place, it is necessary to decide on a method of curve fitting.

Robley Winfrey (44, 45) has not only undertaken to find the proper type of curve to use but he has done the job so thoroughly that, if his methods are accepted, it is unnecessary to go through the labor of making the computations usually required for calculating the constants, or parameters, of the fitted curve. From an extensive study of a variety of different types of industrial property, Winfrey has obtained 18 type curves In use, these curves are drawn on transparent material and laid over the rough curve obtained from the sample One selects from the transparent charts the curve that corresponds most closely and thus obtains not only the type of the curve to be fitted but also its parameters This method should save the labor of curve fitting for all those types of data to which the 18 type curves are applicable

In fact, this work has been done so thoroughly that there is a danger of its being applied to data to which it is not applicable We must be prepared for survivor curves which do not conform to any of Winfrey's 18 types These curves do not, for example, include the case of human mortality, which has two modes, or maxima, while each of Winfrey's curves has only one mode While it appears that industrial mortality is usually unimodal, we should be prepared to face multimodal curves if any ever make their appearance

While it appears that Winfrey's conclusions are correct with regard to the goodness of fit of the various types of curves he tried, these conclusions are apparently based on graphical comparisons and Winfrey's judgment Some scientific test, such as the chi-square test, might have been used to test the hypothesis of goodness of fit in each instance

MULTIPLE DECREMENT METHODS

Property retirements are made for a variety of reasons and a number of causal factors may influence a single retirement In the retirement of a road facility such factors as failure of the foundation may be operative It is important to try to discover, if we can, the reasons for retirements that take place in advance of physical decay of the structure Whether this can be done, the author does not know

At all events, no harm will be done by attempting to distinguish the various causes of retirement, since the figures can always be put together to give total retirement It appears that this work of discrimination among causes of retirement would usually have to be done at the time of retirement rather than at a later date Reasons for retirement could be reported along with salvage values through the use of a procedure such as that proposed by F Thayer Stoddard (31)

UNITS THAT HAVE BEEN USED OR PROPOSED

In any sort of sampling study the selection of a suitable sampling unit is one of the important decisions to be made While the road life studies ordinarily employ sampling methods, in the sense that the investigation is restricted to the few cases of retirement that have been realized, nevertheless the unit employed would be of importance even if the study were thought of as a complete compilation of information Various units have been proposed including the mile of road surface, the square yard of surface, the cubic foot of surface, and the dollar, or a dollar's worth of road improvement Of these units the mile has been most extensively used and the dollar has been used to a limited extent While the use of a square or cubic measure would obviate some of the difficulties inherent in the linear measure, such as varying width and varying thickness of surfaces, it appears that the dollar is for a number of reasons the best of the various units proposed

The main advantage of the dollar as a unit is that it enables the study to take account of salvage values With linear or square measures this would be impossible except in those cases where the usable road surface has been diminished in width or length This might be more frequently practical with a measure of volume, since it would then be possible to take account of diminished thickness of a surface through wear However, deterioration of a structure does not necessarily involve appreciable diminution of its volume, a bridge may be on the verge of falling down and still have the same volume as when new Particularly in the case of foundation work and drainage structures, these units of physical measure appear inadequate While it might be possible to estimate salvage value on the basis of remaining units of service, such as vehicle miles of service, the unit commonly used in appraisals is the dollar

Use of the dollar as a unit has the advantage of immunity from varying purchasing power of the dollar Precisely the contrary has been claimed It has been said that use of the dollar as a unit would necessitate correction for varying purchasing power of the dollar It appears to the author that this charge is based on a superficial analysis It is admitted that the varying purchasing power of the dollar will result in varying quantities of the same type of road construction being purchased for a dollar, but the quantity of construction has no bearing on the life of a dollar invested in it It is no doubt true that the quality of road structure bought for a dollar will also vary with time and that will produce a varying average life, through production of new materials and methods of construction This is one of the things that the road life studies hope to determine It is analogous to the improving mortality rate produced by improved sanitation and methods of medical treatment This constitutes no argument against the use of the dollar as a unit

EXTRAPOLATION OF TIME SERIES

THE PROBLEM OF PREDICTION

In those enterprises, both public and private, which require long-term commitments of capital, the prediction of future values of certain economic time series is of the greatest importance In various other types of enterprises it is sufficient to try to predict for a few months, or at most a year or two, into the future; in highway planning it is important to try to make predictions as much as 20 or 25 years ahead. The attempt to predict so far into the future is beset with extreme difficulties

None of the work that has been done thus far in the planning surveys in an effort to predict future values of time series has any claim to being called scientific. The main excuse for discussing the problem in this paper, which is supposed to be confined to scientific methods, is that the trappings of such predictions may lead someone to think they have a reasonable claim to being called scientific Engineers are accustomed to making estimates. They recognize that their figures are subject to error, and they may estimate roughly the magnitude of the error, even if they do not calculate anything like a probable error or standard error.

When it comes to trying to predict future values of economic time series, the situation is entirely different. The difference is not that the estimate is subject to error but that there is no way of estimating what the magnitude of the error may be. We are no longer in the realm of scientific methods but in the realm of guesswork. This is not to say that such predictions are without value if used with judgment and with a proper appreciation of their unreliability. The author fears that engineers may regard such estimates as being in a class with the estimates they are accustomed to using and treat them accordingly.

CRITICISM OF TRADITIONAL METHODS

In an effort to produce an estimate together with a measure of its reliability, what is frequently done is to fit a trend curve to the data and calculate the dispersion of the deviations of the data from the trend curve. The trend curve is usually a parabola of low degree fitted by the method of least squares. The dispersion of the deviations is supposed to give an indication of the error involved in making predictions from the trend curve. This involves the assumption that the curve will correctly represent the trend of the data in the future and that is a precarious assumption. For one thing, any number of trend curves could be fitted to the data; by using a parabola of sufficiently high degree it would be possible to achieve a perfect fit with the curve going through all the given points. Furthermore, even if we assume that we have a good representation of the trend as it has manifested itself in the past, we have no assurance that the same trend will continue in the future. The trend is liable to change at any moment and there is no way of calculating a probability for this change or the amount of the probable change.

As one looks over past values of those economic series that are directly related to highway planning, such as motor vehicles registered and gasoline consumed, one has the feeling that it would have been extremely easy ten years ago to predict the values up to the present There has usually been a regular sort of growth with little evidence of cyclical variation except for the depression years Even the accidental variations are usually small in size. It may look as though we can draw smooth curves into the future and make as good predictions as we now feel we could have made ten years ago The great difficulty is that we do not know how far into the future this regular growth will continue The human population may tend to become saturated with motor vehicles and motor-vehicle transportation In that event, the trend curves will have a tendency to become level, but no one knows when this leveling tendency will start or how fast it will take place. Inventions, wars, and other unpredictable happenings may change the trend entirely.

It appears to the author that when analysis of time series depends upon assumptions about the shape of curves that we believe nature will conform to, it is in a precarious position. What should be done is to make assumptions about the economic and social machinery at work. For example, a report prepared by the Public Roads Administration indicates that, except for the highest income bracket, around 12 per cent of personal income goes for motor transportation regardless of the income bracket This result, if correct, is of great importance. It has been used in certain instances as a guide in the projection of trend curves into the future. Considerations such as this, rather than assumptions about the shape of the trend curve, should guide extensions of trends into the future

THE OWEN REPORT

The report of Wilfred Owen (27) on "Trends in Highway Financial Practices," presented to the Highway Research Board in December 1939, impresses the author as the best type of approach to this problem that is possible in the present state of knowledge. This report contains no charts, no graphical extensions on semilogarithmic paper. It does not pretend to reduce the making of predictions to an impersonal formula. The assumptions it makes are about the economic and governmental machinery rather than about the shapes of trend curves. This appears to be the correct approach It is in the scientific spirit; wherein it fails of being scientific is in its lack of precision in statement and rigorous demonstration of the conclusions from stated assumptions. Such precision and rigor are not now possible in this kind of work; perhaps they never will be Nevertheless, such discussions may be of great practical value. It is no doubt useful to the economist such as Owen, or to the executive, to have trends of economic time series that have been observed in the past, it is of doubtful value to have them produced into the future because there are so many different ways of producing them, all apparently plausible The plausibility of some one method of extension is liable to mislead the practical man where his judgment would have been a safer guide.

Experimental Methods

THE USE OF EXPERIMENTS

Whatever natural knowledge we have is acquired through experimentation. While at primitive stages knowledge may be so meager that almost any happening brings new knowledge, at higher levels the design of experiments to throw light on definite questions takes on great complexity In fact, the methodology of experimental design has become a branch of study in itself.

While the study of individual instances, short of complete enumeration of all the cases, which is usually impossible, can never suffice to establish the truth of a general proposition, a single example of the failure of a general proposition is all that is needed to disprove it

In the realm of natural knowledge, however, we are usually unable to find even a single example of the failure of a proposition which unequivocally disproves it The most we usually find is an example which is highly improbable if the given proposition is true We may then be led to reject the proposition, but we do not know certainly that it is false; we recognize that we may occasionally reject a proposition that is actually true Nor do we ever accept a proposition; we merely do not reject it and in this sense we may be said to give it a qualified acceptance The most we can say is that we have confidence in the truth of the proposition we do not reject and our confidence is greater the greater the number of critical experiments the proposition is subjected to without being rejected All we mean is that if we act as though the proposition were true we shall get into difficulty less often, usually far less often, than if we act as if the proposition were false

Statistics, rightly understood, is not a magic for squeezing something out of a number of improperly controlled experiments, it is the science of combination of observations when the observations have been made according to the most reliable methods. Not until an experiment has been repeated a number of times by independent investigators do we regard a result as established. Even then, we do not expect the results of repetitions of the experiment to agree, we expect merely that the results will establish a statistical distribution which does not contradict the hypothesis the experiment is supposed to support.

THE DESIGN OF EXPERIMENTS

There are just two qualities of an experiment that are of importance for the practical purposes of the highway planning surveys. An experiment should, foremost, be unbiased, in addition, it is desirable that it be as efficient as possible in relation to its cost By saying that an experiment is unbiased we mean that as the experiment is replicated more and more times, estimates calculated from the experiments approach the correct value more and more closely It is assumed that the estimates calculated from the experimental data are correctly calculated. In fact, the design of an experiment should include methods of calculating statistics which will provide unbiased and efficient estimates of the parameters which it is desired to estimate

By saying that an experiment should be efficient we mean that the dispersion of the errors should be a minimum The amount of information yielded by an experiment obviously increases as the dispersion of the error of the experiment decreases It is convenient in practical work to use a unit of information introduced by R A Fisher (10). An experiment which produces a standard error of the mean of 1 percent is said to yield 100 units of information; a standard error of the mean of 5 percent yields 4 units of information, a standard error of the mean of 10 percent yields 1 unit of information. Thus the amount of information is regarded as varying inversely with the variance of the mean.

An example of the most efficient design of an experiment is found in the method of representative stratified sampling, discussed in the next section. The most efficient allocation of the sample to the strata is secured by having it proportional to the product of the number of elements in each stratum of the population by the standard deviation of that stratum If the sample is allocated to the strata in proportion to the number of elements of the population in each stratum, then the design is still unbiased but it has lost its property of being a most efficient design. Furthermore, the system of stratification that is most efficient is the system that minimizes the intrastratum variance.

A complete design of an experiment includes unique statistics which are both unbiased and efficient In practice, however, sampling procedure is often devised by one person or organization while the devising of methods of statistical analysis is left to another person or organization, possibly even after the information has been collected. It may be found that the sampling method was so poorly designed that unbiased statistics are out of the question. But even if the experiment has been properly designed, it is possible to waste much of the information provided by the sample through the use of inefficient statistics. As an example, suppose that a sample has been collected in accordance with the method of representative stratified sampling. An incompetent statistician might calculate the simple arithmetic mean instead of the weighted mean required by that method of sampling The estimate so provided would be unbiased but it would not be the most efficient estimate; it would waste part of the information provided by the sample. Such an estimate would have a larger standard error than the proper weighted average.

Sampling Methods and Their Statistics

RANDOM SAMPLING

If what we have to sample from is a statistical population, and if we have no knowledge which would make possible some type of stratified sampling, then the best that can be done is to draw a sample at random from the population. Even in stratified sampling we are expected to draw a sample at random from each stratum. The importance of random sampling thus makes it fundamental in any discussion of sampling It must be emphasized that random sampling is applicable in case the thing we are sampling from is a statistical population. Ordinarily in practice we do not have a definite statistical population to sample from, what we usually have is a series of population arranged chronologically or geographically in some kind of order. Thus our problem is usually one of quadrature rather than one of sampling from a population. Confusion of these two problems has led to many curious arguments in statistical literature.

A sample is said to be drawn at random provided the probability of a given element of the population being in the sample is constant for all elements of the population. Suppose, for example, that we have a perfectly balanced roulette wheel with 10 digits, then when we spin the wheel the probability of a given digit is 1/10. Or suppose that we thoroughly shuffle a deck of cards and draw out a hand of 5; the probability that any particular card is in the hand drawn is 5/52. Of course any hand that is actually drawn, no matter how much the cards were shuffled, is an improbable hand, but the point is that with infinitely many repetitions of the process each card will be drawn the same number of times as any other.

If the elements of our population are numbers and we calculate the mean of the numbers drawn in a sample, we shall find that different samples frequently produce different means. For a few populations the distribution of the means from random samples has been calculated mathematically. This is notably true for the normal distribution. In most cases, however, either we do not know the analytical form of the distribution we are sampling from or the mathematical problem of obtaining the distribution of the sample means is too difficult for us to solve. The usual practice now is to assume that the population we are drawing from is normal and to use the table of the normal distribution. It is, of course, true that the normalized sum of n independent random variables of equal dispersion tends to the normal distribution as n increases indefinitely. This theorem brings it about that the mean of a large random sample, even if drawn from a non-normal population frequently tends to be normally distributed. It should be remembered that a very large sample may be required for a close approximation to the normal distribution.

STRATIFIED SAMPLING

Proportional and Representative Stratified Sampling

The method of stratified sampling has been developed during the past few years by A. L. Bowley (2), J. Neyman (22), and other statisticians The early form of this method, introduced by Bowley, prescribed drawing a sample of the same proportion from each stratum. The efficiency of the method has been improved through the work of Neyman who has determined the most efficient allocation of the sample among the strata. We may call Bowley's method the method of proportional stratified sampling while Neyman's method will be called the method of representative stratified sampling. Neyman's method is also known as the representative method or simply the method of stratified sampling. The method of proportional stratified sampling is the most efficient for certain types of investigations such as public opinion polls requiring yes or no answers.

It appears that the method of representative stratified sampling is the best that is known at the present time. Various questions relating to its application present themselves, however. This method involves a combination of systematization and randomization that makes the most effective use of available knowledge. A statistician cannot be said to have done his full duty unless he makes use of all available, pertinent knowledge in the design of the sampling procedure he employs in a particular study. Available knowledge and practical ways of collecting the information always condition the sampling procedure; this conditioning manifests itself in the manner of applying the method of representative sampling.

System of Stratification

The object of stratification is to divide the population into subpopulations which are internally as homogeneous as possible and which are as heterogeneous among themselves as possible. These are just the contrary of the qualifications that a good sampling unit should possess. However, if we consider what our real object is, there is no antithesis In any population sampled we prefer to have the members of the population as much alike as possible. If they are all precisely similar in the characteristics we are interested in, then we have to draw only one member in order to get a perfect representation of the population. In stratifying a population our object is to divide it into subpopulations each of which will be internally as homogeneous as possible. It follows that the heterogeneity of the original population must to a considerable extent be reflected in differences among the subpopulations.

For the system of stratification adopted it is necessary that we have available a multivariate table showing the simultaneous distribution of the population with respect to the set of controls used in the stratification. This multivariate distribution is necessary both in order to determine the proper number of samples to draw from each stratum and also to make possible the expansion of the sample data after the sample is drawn. If, for example, we have decided to use vehicle type, county of registration, year model, and make of vehicle as controls, then we need tables, such as those prepared by the Reuben H. Donnelley Corporation, which show the simultaneous distribution of these variates.

Usually practical considerations, particularly the economy of collecting the data, must govern the stratification However, out of those systems of stratification that are available it may be possible to select one that is superior to the others. One system of stratification is better than another provided the intrastratum variance of the first is less than the intrastratum variance of the second. It is assumed that the variance is calculated for the variate we are interested in or the most important of the variates we are interested in

Special Problems in Application of Representative Sampling

Formulas for the most efficient allocation of the sample among the strata have been derived by J Neyman (22). These formulas take account of the size of each stratum as well as its variability, or dispersion. The greater the size of the stratum of the population, other things being equal, the greater the size of the sample that should be drawn Also, the greater the variability, or heterogeneity, of the stratum, other things being equal, the greater the size of the sample that should be drawn. If the dispersion of the variate we are interested in is constant for all the strata, then we have nothing but the sizes of the strata to consider and we draw a sample that contains a uniform proportion of each stratum A statement of Neyman's formula that is very nearly precise is that the sample should be allocated to the strata in proportion to the product of the number of population elements in each stratum by the standard deviation of that stratum.

The method of representative stratified sampling involves unique statistics for the estimation of the mean and the standard error of the mean. Formulas for these two statistics have also been derived by Neyman (22). The procedure is very similar to that used in the case of a sample drawn at random from normal population. The author (37, 38) has devised punched card technique for the more rapid and economic calculation of the estimate of a mean by Neyman's formula.

METHOD OF DOUBLE SAMPLING

It sometimes happens, in attempting to apply the method of representative sampling, that a proposed system of stratification would be highly efficient but cannot be used because information about the distribution of the proposed control variates in the population is not available Provided the use of the proposed system of stratification would effect enough improvement in the reliability of the desired statistics, then it may be justifiable to obtain the needed control information by means of a preliminary random sample. This leads to what has been called double sampling by Neyman (24). Special formulas relating to double sampling have been developed by Neyman to show the proper allocation of available funds between the preliminary random survey and the final representative survey

An example of the utility of double sampling is presented by the motor vehicle allocation and road use studies The motor vehicle allocation study undertook to determine the distribution of the motor vehicles by population groups and the miles of travel of each vehicle for the past year In Texas the distribution of vehicles by population groups given by M V A was utilized as control information in drawing the road use one percent sample. It has been suggested that total miles of travel would probably have furnished an excellent control and this seems entirely plausible No attempt was made, however, to secure the most efficient allocation of available funds between the two surveys, or indeed to determine whether the M.V.A. study was justified at all. Neyman's formulas were not published until March 1938, too late to be of any assistance in the initial survey. These formulas, however, should be of great value in future studies of this type.

ORGANIZATION OF STATISTICAL WORK

It appears to the author that highway planning survey work has suffered from lack of recognition of the important place of statistical work, in the general sense of that term In most instances statistical work appears to be an afterthought and an adjunct in the planning surveys. A statistician, if employed at all, is expected to calculate statistics from data already collected, rather than to design unbiased and efficient methods for collecting information, which is his main function. This situation is particularly surprising and deplorable, since the planning surveys represent one of the largest statistical projects ever undertaken.

Perhaps the attitude of statisticians themselves is in considerable measure responsible. Statisticians have tended to place too much emphasis on statistics and to neglect the study of the design of sampling methods There has been a disposition, too, to regard statistics as a branch of economics So long as this is done it is understandable that statistics will not command the respect of natural scientists and engineers Properly understood, statistics is a branch of applied mathematics, that branch which applies mathematics to experimental technique, or to observational data Statistics has many points of analogy with that branch of mechanics called statics; it has analogies also with the theory of surface zonal harmonics used by electrical engineers.

The proper practice of statistics, including the design of methods of collecting data, requires an intimate knowledge of the field in which statistics is being applied. A good biometrician must have a close acquaintance with biological studies; an actuary must make a special study of problems of life contingencies Similarly, a highway statistician needs a close acquaintance with highway work. It would no doubt be better for a highway engineer to study statistics than for an economist to undertake to study engineering. For the engineer, statistics would simply be another subject, like mechanics, for most economists the study of engineering would require the mastery of a whole new field for which they might have little propensity.

The problem is made more complicated by the primitive state of development of the subject of statistics. The subject is largely in its infancy and methods adequate for the problems that arise in practice usually do not exist It is frequently necessary for the practical statistician to solve new mathematical problems or to rely upon his intuition when he is unable to produce the solution in a rigorous manner. In a subject that relies so heavily upon intuition as statistics does in its present state of development, the mathematical competence of the statistician is of prime importance Statistics is not yet a subject to get a professional degree in and practice from a handbook

The author's conclusion is that one of the most beneficial things that could befall highway planning survey work would be for the engineers engaged in this work to study statistics No doubt some of them have, the work of Winfrey and other engineers gives evidence of that. The simplest thing to do would be to bring in young graduates in statistics from the universities In addition to the consideration that their training is usually far from adequate, there is the further consideration that they would have to learn highway engineering. A better proposal might be to bring in college professors and young instructors in statistics However, thorough understanding of mathematical principles does not necessarily insure skill in their application and years spent in abstract research frequently dispose one unfavorably to practical work

Whatever may be the best organization of government statistical work as a whole, it appears clear to the author that planning survey work would benefit from much closer centralization. Methods of sampling and statistical analysis should be adopted by professional statisticians of the highest rank. These methods, once adopted, should be adhered to closely by all organizations engaged in the work Such a system need not stifle individual initiative Individuals anywhere in the organization who are competent to do so should be free to make suggestions regarding the improvement of methods; they should not be free to apply their suggestions in the conduct of the work until they have been thoroughly considered and tested experimentally.

It is important to make a clear distinction between statistics and propaganda All that is required of a statistic is that it be unbiased and that its random error be as small as possible Honesty alone does not insure these characteristics, it requires the utmost care to produce them It is easy enough for enumerators to bias replies unconsciously The road use enumerators, for example, are liable to bias the replies favorably to the State highway system because of their announcing themselves as representatives of the State highway department and because the emphasis in the headings of the forms, in the maps, and throughout the discussion is on the State highway system and the State highway department It appears that any administrator would have as his first concern knowing the facts about his business and his situation, however unpleasant those facts might be, and however carefully he might guard their secrecy.

REFERENCES

1 Dinsmore Alter, "A Simple Form of Periodogram," The Annals of Mathematical Statistics, Vol 8 (1937), pp. 121-126

2 A L Bowley, "Measurement of the Precision Attained in Sampling," Bulletin, International Statistical Institute, Vol 22, Book 1

3 Nathan Cherniack, "Methods of Estimating Vehicular Traffic Volume with the Aid of Traffic Patterns," *Proceedings* Highway Research Board, Vol 16 (1936), pp. 253-269

4 R. E. Craig, "An Automatic Recorder for Counting Highway Traffic," *Public Roads*, Vol 19 (1938), pp 37-51.

5 R E Craig and S. E Reymer, "A Simple Portable Automatic Traffic Counter," *Public Roads*, Vol 19 (1939), pp 213-221

6 Harald Cramer, "Random Variables and Probability Distributions," Cambridge Tracts in Mathematics and Mathematical Physics No. 36, University Press, Cambridge (1937).

7 Edward L. Dodd, "Periodogram Analysis with the Phase a Chance Variable," *Econometrica*, Vol. 7 (1938), pp 57-63

8 Arne Fisher, "Mathematical Theory of Probabilities and Its Application to Frequency Curves and Statistical Methods," New York (1936)

9 Arne Fisher, "Statistical Hypotheses and Estimation in Historical and Critical Light," presented to the Cowles Commission Research Conference on Economics and Statistics (1938)

10 R A Fisher, "The Design of Experiments," London (1937).

11 R A. Fisher, "Statistical Methods for Research Workers," 6th ed, London (1936).

11a C H French, "Study of Reliability and Sampling Methods of the Texas Road Use Survey," Public Roads Administration, Washington, D C (1941) (mimeographed).

12 C F Gauss, "Theoria Combinationis Observatorium Erroribus Minimis Obnoxiae," Goettingen (1821)

13 L R Hafstad, "On the Bartels Technique for Time-Series Analysis," *Journal*, American Statistical Association, Vol 35 (1940), pp 347-361

14 E H Holmes, "Procedure Employed in Analyzing Passing Practices of Motor Vehicles," *Public Roads*, Vol 19 (1939), pp 209-212, 221

15 E. H Holmes and S E Reymer, "New Technique in Traffic Behavior Studies," *Public Roads*, Vol 21 (1940), pp. 29-45

16 W R Houston, "The Art of Treatment," New York (1936)

17 Sigvald Johannessen, "Highway Economics," New York (1931)

18 M G Kendall and B Babington Smith, "Tables of Random Sampling Numbers," Tracts for Computers, No 24, Cambridge University Press (1939).

19 Edwin B Kurtz, "Life Expectancy of Physical Property," New York (1930).

20 F. C Mills, "Statistical Methods Applied to Economics and Business," New York (1935)

21. L M Milne-Thomson, "The Calculus of Finite Differences," London (1933)

22 J. Neyman, "On Two Different Aspects of the Representative Method," *Journal*, Royal Statistical Society, Vol 97 (1934), pp 558-625

23 J. Neyman, "Lectures and Conferences on Mathematical Statistics," Graduate School, U S Department of Agriculture, Washington, D C (1938).

24 J Neyman, "Contribution to the Theory of Sampling Human Populations," *Journal*, American Statistical Association, vol. 33 (1938), pp. 101-116

25. O. K. Normann, "A Simple Accumulating-Type Traffic Counter," *Public Roads*, Vol 20 (1939), pp. 195-196, 203.

26 O. K. Normann, "Preliminary Results of Highway Capacity Studies," *Public Roads*, Vol 19 (1939), pp 225-232, 240

27. Wilfred Owen, "Trends in Highway Financial Practices," *Proceedings*, Highway Research Board, Vol 19, p 15 (1939).

28 Max Sasuly, "Trend Analysis of Statistics, Theory and Technique," Washington, D C (1934).

29 W Arthur Shelton, "Mehods of Estimating Highway Traffic Volume," *Proceedings*, Highway Research Board, Vol. 16, p 239 (1936).
30 Eugen Slutzky, "The Summation of Random Causes as a Source of Cyclic Processes," *Econometrica*, Vol 5 (1937), pp 105-146
31 F Thayer Stoddard, "Suggested Plan for Continuing the Salvage Value Study," Austin, Texas, January 15, 1940 (typewritten)
32 R O Swain, "A Practical Application of Observing Traffic by Means of Extremely Short Counts," Austin, Texas (1939) (mimeographed). Abstract published in *Traffic Engineering*, Vol 9, No 18 (October, 1939)
33 R O Swain, "Traffic Aids to Texas Municipalities," *The American City*, Vol. 55, No 7 (July, 1940), pp 85-89
34 L H C Tippett, "Random Sampling Numbers," Tracts for Computers No. 15, University Press, Cambridge (1927).
35 C W Vickery and V L. Beavers, "Extensional Blanket Analysis and Field Methods," Austin, Texas (1937) (Mimeographed).
36 C W Vickery, "The Construction of Traffic Flow Maps and the Determination of Average Trip Length from Road Use Data," Austin, Texas (1937) (Mimeographed).
37 C W Vickery, "Punched Card Technique for the Correction of Bias in Sampling," *Journal*, American Statistical Association, Vol 33 (1938), pp 552-556
38 C W. Vickery, "Note on Punched Card Technique," *Journal*, American Statistical Association, Vol 34 (1939), pp 110-111.
39 C W. Vickery, "On Drawing a Random Sample from a Set of Punched Cards," Supplement to the *Journal*, Royal Statistical Society, Vol 6 (1939), pp 62-66
40 C W Vickery, "Deformations of Frequency Functions," presented to the American Mathematical Society, April 8, 1939 Abstract published, *Bulletin*, American Mathematical Society, Vol 45 (1939), p 378
41 C W. Vickery, "Random and Biased Sampling Machines," presented to the American Mathematical Society, April 8, 1939 Abstract published, *Bulletin*, American Mathematical Society, Vol. 45 (1939), p 378
42 C W Vickery, "On Estimating Street Traffic by Means of Extremely Short Counts," Austin, Texas, 1939 (Mimeographed) Abstract published in *Traffic Engineering*, Vol 9, No 17 (September, 1939), pp 5-6
43 C W Vickery, "Estimating Street Traffic by Extremely Short Counts," *Highway Research Abstracts* No 65 (November, 1939), pp 8-9.
44. Robley Winfrey and Edwin B Kurtz, "Life Characteristics of Physical Property," Iowa Engineering Experiment Station Bulletin 103, Ames, Iowa (1931).
45 Robley Winfrey, "Statistical Analyses of Industrial Property Retirements," Bulletin 125, Iowa Engineering Experiment Station, Ames, Iowa (1935)
46. "Depreciation Charges of Telephone Companies," American Telephone and Telegraph Company, New York (1928)
47 "Government Statistics," a report of the Committee on Government Statistics and Information Services, sponsored by the American Statistical Association and the Social Science Research Council, New York (1937)