SAMPLE-SIZE DETERMINATION

Arthur E. Mace
Battelle Memorial Institute
Columbus, Ohio

Reinhold Publishing Corporation, *New York*
Chapman & Hall, Ltd., London

Library of Congress Catalog Card Number: 64-18673
Printed in the United States of America

Preface

In planning experiments, engineers and scientists are invariably confronted with the question: "How many observations should be made?" This engineering manual is designed (1) to familiarize experimentalists with the kind of information required for sample-size determination, and (2) to provide them with elementary procedures, formulas, and tables for choosing an economical sample size.

Essentially this manual is a compilation of mathematical procedures for determining the optimum size of a research experiment, that is, the number of specimens which should be observed. Some forty different types of objectives for a research experiment are considered. For each type of research objective, a mathematical formula is developed to determine the minimum number of observations necessary to achieve that objective. A set of twenty tables of common statistical distributions is provided in the appendix to assist in applying the formulas to actual problems in the planning of research experiments. For each type of research objective, a case example of applying the formulas in the design of experiments in electronic reliability has been carried out. Additional information is included on lot acceptance sampling plans of the U.S. Department of Defense and on methods for constructing cost schedules for sample-size determination.

The first draft of this manual was written in 1960–1961 as a guidance document for the laboratory evaluation of electronic component parts under the sponsorship of the Electronic Component Reliability Center at Battelle Memorial Institute. We are indebted to the following organizations which have granted permission to publish this material: Bell Telephone Laboratories; the Boeing

Company; General Dynamics Corporation; General Electric Company; General Precision, Inc.; Hughes Aircraft Company; International Business Machine Corporation; Lockheed Aircraft Corporation; Martin-Marietta Corporation; Motorola, Inc.; National Aeronautics and Space Administration; Radio Corporation of America; the Raytheon Company; and Westinghouse Electric Corporation.

Subsequently helpful comments for revision of the manual were received from several representatives of these sponsoring organizations, from Professors Robert Bechhofer of Cornell University and K. A. Brownlee of the University of Chicago, and from my colleagues at Battelle, G. H. Beatty, C. W. Hamilton, and R. E. Thomas. Professor Charles Quesenberry of Montana State College read the first draft in detail during the summer of 1962 and in the course of our many discussions contributed much to the revision.

For the tables which appear in the appendix I am indebted to E. S. Pearson, editor of *Biometrika;* Clifford Hildreth, editor of the *Journal of the American Statistical Association;* and J. L. Hodges, Jr., editor of the *Annals of Mathematical Statistics,* for permission to reprint tables which have appeared in their journals. I am also indebted to E. S. Pearson and H. O. Hartley and to the publisher, Cambridge University Press, for permission to reprint Tables 12 and 18 from their book "Biometrika Tables for Statisticians," Volume I; to Sir Ronald Fisher and Frank Yates and to the publisher, Messrs. Oliver and Boyd, for permission to use numerical entries in Tables III and IV from their book "Statistical Tables for Biological, Agricultural and Medical Research"; and to Anders Hald and to the publisher, John Wiley and Sons, for permission to reprint Table XII from his book "Statistical Tables and Formulas."

Finally, I express my appreciation to Battelle Memorial Institute for encouragement and support in this effort and to Mrs. Nancy Zacaroli, secretary of the Mathematics and Theoretical Physics Division, for the difficult typing of both the first and the final drafts of the manuscript.

ARTHUR E. MACE

April 1964

Contents

CHAPTER 1

The Nature of Statistical Inference

Modern engineering relies heavily on sampling experiments as a source of information for making engineering decisions. Chemical companies test small laboratory batches of a new paint formula to estimate the durability of the paint in actual service. Automobile manufacturers inspect samples of incoming shipments of bolts and nuts to compare lot quality with acceptance specifications. Computer manufacturers life-test small samples of electronic parts from each of several vendors to select the best source of supply for use in new circuit designs. Aircraft materials laboratories fatigue-test specimens of new metals with different combinations of chemical elements in order to compare the effects of these elements on tensile strength.

1. Inferences and Their Accuracy

The results of such sampling experiments are used to predict the consequences of making specific engineering decisions. From the observed behavior of a relatively small number of observational units, inferences are made about the unobserved behavior of usually a much larger number of units. Two sources of error can contribute to the inaccuracy of such inferences:

(1) *Errors of bias*, which arise principally because the particular conditions of the experiment almost never represent the exact model of the situation about which inferences are desired.

(2) *Errors of sampling*, which arise principally because the particular observational units chosen for the sample almost never represent the exact population of units about which inferences are desired.

An experiment can be designed to minimize the principal source of biased error by specifying the conditions of the experiment so that they are physically similar to the situation about which inferences are desired. There are a number of excellent accounts of the theory of modeling and dimensional analysis available for guidance in formulating such specifications.*

An experiment can also be designed to control the principal source of sampling error by specifying the number of observational units to be drawn at random from the population of units about which inferences are desired. This manual is designed to provide computational formulas for use in specifying the sample size.

2. Some Definitions of Terms

The mathematical justification for making statistical inferences depends upon what is known as random sampling. To describe this relation, we begin by defining some terms:

Observational Unit. An individual member of a collection of nominally identical objects or responses of objects to treatment, identifiable by one or more variables. It is presumed that the individual identification of the member within the collection arises from sources of variation which are independent of those affecting any other individual member.

Population. A finite or infinite and sometimes hypothetical collection of nominally identical observational units about which inferences are desired.

Random Sample. A finite subset of observational units drawn from a population by a chance process. It is presumed that on each draw all units in the population have had an equal chance of being selected for the subset.

Observation. One or more variables which describe an individual observational unit included in the sample, either directly observed or derived from measurements.

* Langhaar, H. L., "Dimensional Analysis and Theory of Models," New York, John Wiley & Sons, 1951; Birkhoff, G., "Hydrodynamics. A Study of Logic, Fact, and Similitude," Princeton, Princeton University Press, 1950; and Bridgeman, P. W., "Dimensional Analysis," New Haven, Yale University Press, 1949.

Sample Statistic. A summary variable, derived from observations, which describes either the distribution of observations in one sample or the configuration of such distributions in two or more samples, frequently used to estimate some population parameter.

Population Parameter. A constant of unknown value which describes the distribution of variables representing observational units in one population or the configuration of such distributions in two or more populations, frequently estimated by some sample statistic.

3. The Basis for Inferences

The objective of the sampling experiment is to make unbiased and reasonably precise inferences with respect to the consequences of engineering decisions. The particular engineering decision should specify (1) the observational unit involved in the decision, (2) each population affected by the decision, and (3) the population parameter which measures the consequence of the decision. Then the basic problems to be solved in designing the sampling experiment are to determine (1) the sample statistic which will be used to estimate the population parameter, (2) the sampling process for selecting observational units so that a random sample of observations from the population can be expected, and (3) the minimum number of observational units needed to furnish inferences of desired precision.

Mathematicians have devised procedures for making inferences about population parameters from a set of observations provided these observations constitute a random sample of independent identically distributed random variables. To the extent that engineers and statisticians are able to specify a sampling process which should produce such a random sample, there is reason to expect that the set of observations actually produced can be used in making unbiased inferences about population parameters.

Mathematicians have further studied the probability distributions of certain sample statistics as a function of the number of observations in a random sample used in their derivation. As the sample size increases, sample statistics derived from observations tend to have smaller sampling errors. To the extent that engineers and statisticians are able to prescribe the number of observational units for controlling sampling errors, there is reason to expect that the sample statistic actually produced can be used in making inferences of arbitrarily predetermined precision.

4. Types of Inferences

There are three broad categories of problems for which systematic procedures have been developed to exercise control over sampling errors:

(1) *Estimation Problems.* Problems in which we wish to infer that the true but unknown value of a specified population parameter is contained within a bounded interval of given width.

(2) *Tests of Hypotheses.* Problems in which we wish to make one of two inferences: either that the true but unknown value of a specified population parameter differs from a specified standard in one or both directions or that any difference is less than a given amount.

(3) *Selection Problems.* Problems in which we wish to select from several populations the one with the highest (or lowest) true but unknown value of a specified population parameter when the assumed difference between the highest and next highest values (or between the lowest and next lowest values) among the various populations is at least a given amount.

The procedure for making these statistical inferences requires that prior to the conduct of the experiment the engineer or statistician assign relatively low probability levels, say 10 per cent, to the values of the maximum risks of making an incorrect inference he is willing to take. Then after the experiment has been performed, and assuming the observations produced constitute a random sample of independent identically distributed random variables, the engineer or statistician can make unbiased inferences about population parameters with preassigned maximum risks of being wrong.

5. The Precision of Inferences

With fixed probability levels assigned to the maximum risks of making an erroneous inference, the way in which the precision of the inference can be measured depends upon the particular category of problem:

(1) *Estimation Problems.* The precision of the inference can be measured by the width of the confidence interval.

(2) *Tests of Hypotheses.* The precision of the inference can be measured by the width of an indifference margin—the maximum difference between the true but unknown value of the population parameter and the standard which is to be allowed without limiting to a preassigned probability level the risk of failing to detect this difference.

(3) *Selection Problems.* The precision of the inference can be measured by the width of an indifference margin—the maximum difference between the highest and the next highest true but unknown values of a population parameter (or between the lowest and next lowest values) among several populations which is to be allowed without limiting to a preassigned probability level the risk of selecting an incorrect population as the one with the highest (or lowest) population parameter.

Once a sampling experiment is completed, the engineer or statistician can compute the width of the appropriate confidence interval or indifference margin based on the observed values. The precision with which inferences can be stated turns out to be a function not only of the number of sample observations but also of the variability in the values of these observations. Large variability in observed values and/or small sample sizes produce imprecise inferences. Small variability in observed values and/or large sample sizes produce precise inferences.

When a sampling experiment is designed for purposes of making engineering decisions, there are two kinds of information required to determine the minimum sample size so that inferences of desired precision can be expected. The first kind of information required is a statement of the experimental objective which should contain a specification of the amount of precision in inferences about the designated population parameter which is of practical significance for decision. The second kind of information required is a quantitative statement of the variability in observed values which is to be expected. The next two chapters indicate how each of these two kinds of information can be obtained for purposes of determining an economical sample size.

CHAPTER 2

The Objective of a Sampling Experiment

The design of economical sampling experiments for purposes of making engineering decisions requires a clear understanding of the experimental objective. If the individual responsible for experimental design does not have this understanding, the experiment is doomed to failure even before it begins. On the other hand, when he does have this understanding, there is reason to expect that requirements for data analysis can be correctly anticipated at the time the experiment is designed.

In modern engineering the same individual is seldom completely responsible for the design of the experiment, the conduct of the experiment, the analysis of the experimental data, and the making of the engineering decision. With a group of different individuals involved in the process of experimentation and decision-making, it is essential that a written statement of the experimental objective be prepared. Experience indicates that this statement should contain specific and concise operational definitions of the following:

(1) the observational unit involved in the decision,

(2) the one or more populations involved in the decision,

(3) the population parameter which is to provide the basis for decision,

(4) the type of inference about the population parameter required for decision, and

(5) the amount of precision in inferences about the population parameter which is of practical significance for decision.

1. The Importance of Operational Definitions

The purpose of these definitions is to direct the actual physical process of sampling, experimentation, and observation toward the goal of producing plausible inferences relevant to making particular engineering decisions. For any particular experiment these definitions will be determined largely by the subject matter of the engineering decisions to be made. But there are important design factors which should be considered so that these definitions will strengthen the plausibility of any inference made after the sampling experiment is completed.

By the plausibility of an inference we mean the extent to which assumptions about the hypothetical process producing sample observations and justifying inferences are in fact satisfied by the physical processes actually used. The procedure for making inferences implies that after a finite number of sample observations are completed, probability statements will be made as to what the results of the experiment would be like if the number of observations were not terminated but allowed to approach infinity. The procedure assumes that any set of sample observations produced by the experiment will constitute a random sample of identically distributed random variables.

Whether any actual physical process will produce such a set of sample observations is not known, and whether any actual set of sample observations is in fact such a set of sample observations cannot be determined with certainty. The plain fact is that, for purposes of designing engineering experiments and making inferences based on sample observations, there exists a mathematical model for making valid inferences from sample observations on the assumption that these sample observations have been produced by a hypothetical process of sampling, experimentation, and observation. To the extent that the engineer or applied statistician can direct an actual physical process of producing sample observations toward this hypothetical process by defining the conditions for sampling, experimentation, and observation, he should be able to strengthen the plausibility of any inference that he may wish to make.