Broadening the Horizons of Experimental Design

LT GEN William B. Bunker
U.S. Army Materiel Command
Washington, D.C.

From its early beginnings, statistics has been an important vehicle with which reasonable men have attempted to seek an understanding of the problems which confront them. Some of the earliest developments and applications of statistical concepts occurred in response to problems at the gaming tables. In fact, I have been told that more than one early statistician earned his keep by calculating odds for a wealthy gambler. The basic orientation of statistics toward the solution of practical problems can be found as the motivation for many major developments in statistics. For example, Thomas Bayes in his often quoted and controversial essay stressed his desire to provide a more efficient procedure for the estimation of probabilities. More recently, the contributions of Professor R.A. Fisher in the area of small sample statistics were motivated by a desire to improve the analytic tools available in biomedical research.

The essential point is that many of the important developments in statistics were motivated by a desire to solve real world problems. I am concerned that in some quarters this orientation to problem-solving has been replaced with a tendency toward self contemplation and a primary interest in statistical purity. There is a need to re-examine the direction of current efforts and to confront our major problems head-on. Only through broadening the horizons of experimental design can we hope to deal effectively with our most pressing problems.

Today, as a first step toward broadening the horizon, I would like to spend the remainder of my time discussing several areas that are amenable to the application of the concepts of experimental statistics.

SYSTEM TESTING AND DEVELOPMENT

One important area in which much work is needed involves the statistical issues in equipment testing. At the outset, I want to stress that our test programs are not and in fact cannot be scientific experiments. One reason for this is the traditional requirements for the design of experiments are infeasible within the context of a test and development program. For example, a basic principle of design of experiments involves the control or minimization of the variation in the experimental situation. This is an almost impossible requirement to satisfy for two reasons. First, due to modification in the system during development, the basic heterogeneity of experimental units is high. This inherent variability represents a violation of a basic statistical assumption. Second, the dimensions of the problem frequently preclude control or even measurement of extraneous sources of variation. The problem was illustrated in the test program for our new AAFSS.

The status of a scientific experiment also is denied to our development and test programs because of the fact that we just can't afford the large number of data points that are required in a classical experimental design. In practice, testing is done on a small number of prototype systems. If an attempt was made to gather the number of observations required to achieve the desired level of statistical significances, no development would ever take place.

The statistical aspects of testing programs are further compounded by our difficulties in specification of the model. In many of our test programs it is difficult to begin to select the relevant variables and logically impossible to identify the important interactions and nonlinearities.

Our recent experience with the development of 152 ammunition for the Sheridan provides a case in point. The variable of interest in this case is binominal, either the round fires or it does not fire. We know that reliability of this ammunition is a function of a number of variables including quality control, the efficiency scavenger system, the ammunition case, and the storage environment, but we also realize that there are n other important dimensions of the problem which remain to be identified. For example, through observation we have established an interaction between the degree of moisture in the powder and the quantity of residue. Experience has demonstrated that higher moisture content resulted in more residue. In response to this finding we have lowered the moisture content, but this change raises a question concerning other yet unknown interactions that are at work in determining the reliability of the ammunition.

Changing the moisture content also illustrates another problem that pervades the testing programs. When the nature of an item is altered as a matter of course in testing and development, how does one aggregate the test data that were generated prior to the change with that data which have been gathered after the change? In a strict sense, the modification has changed the basic structure of the situation that is being modeled, and has made the two sets of data incommensurable. In reality, we are measuring a series of separate probability curves and are reporting the envelope of these curves. This is analogous to developing a baseball batting average by combining performance in the preliminary grapefruit league with that in standard league play. In both cases, the cumulative measure of performance combines early and tentative results with those that have been obtained after the system has been brought up to working order. The net effect of this procedure is to substantially understate the reliability of the system.

Given this situation, how can we give our customer a valid statement of quality assurance? Upon examining the results of the testing program, the statistician would say that we have a ratio of approximately 1 to 52,000, but what we really need to satisfy the customer is a ration of 1 – 1,000,000. At this point I can say, qualitatively, that the real reliability of the system is understated; however, it is impossible to specify the absolute magnitude of the error. Naturally, the customer is not satisfied with the statement about reliability of the ammunition, and something must be done to improve the situation. The statisticians' answer to this dilemma is more testing to develop the required observations. This is an extremely costly procedure and it would have been better to have done more work on estimating the initial function. Ad hoc testing at this juncture is not a feasible solution to the problem.

An alternative approach can be found in the area of statistical decision theory. Resolution of this dilemma may be achieved through the combination of the subjective judgment of the experts and objective experimental results.

A second area in our testing program that requires attention involves the development of large, expensive systems. The Main Battle Tank provides a good illustration of the problem. We really have only a vision of the MBT. In this situation, the problem is that there is no real testing of the whole system. Instead, tests are conducted on different vehicles with various configurations. This means that most of the parameters of interest vary from test to test and that very little remains constant among the tests. What we are attempting to model then is really a function of functions. Casual factors can no longer be expressed simply as numeric values but themselves must be represented as functions, the values of which are in turn dependent upon the value of the total function.

One analytical technique that has been utilized to attempt to model a function of functions is dynamic programming. In the development of the basic algorithm Bellman used a recursive scheme to reflect the method of sequential calculation that is the essence of the approach. For example, consider an aerial weapon system consisting of a navigational subsystem, a target acquisition subsystem, and a weapon subsystem. It is desired to determine the optimal characteristics of all three subsystems, but all these decisions are interdependent. The thing we do know is that whatever navigational and target acquisition subsystems are chosen, the characteristics of the weapon system, e.g., the rate of fire must be optimal with respect to the effectiveness of the whole system. Using the principle of optimality proposed by Bellman, we can say that the optimum rate of fire is a function of the effectiveness of the aerial weapons system. Since we do not know the optimal characteristics for the other two subsystems, the optimal rate of fire and total system effectiveness must be found for all feasible outputs of the subsystem. This technique may provide a clue regarding the way to handle complex equations without knowing their specific form.

The essential point is that we must move away from concepts that require the testing of a static system. Pressures imposed by necessary modifications of systems in the development process do not allow all other things to remain equal and this dynamic aspect of the environment cannot be ignored.

On balance, it appears that increased emphasis on rigor in the design of experiments has diverted our attention from the ultimate objectives. Efforts must be undertaken to develop techniques which provide feasible solutions to problems of quality assurance and the manipulation of more complex dynamic models. We need to soften the science of experimental design to make it a more useful tool in test and development programs. The alternative to this change is to continue to strive for more technically precise answers which are even less meaningful in the decision making process. Unless a conscious effort is made to avoid this plight, experimental statistics may create a paradox similar to that caused by managerial accounting. As a tool of management, the discipline of accounting has experienced an increase in the precision with which financial information is analyzed and reported, but it still does not provide much assistance in the decision making process. Decision makers can safely rely on accounting to identify the loss after the investment has failed, but it is of no help in forecasting the likelihood of this occurrence. It is an after the fact discipline, and our requirements are for knowledge before the fact.

While reflecting on these challenges that lie ahead, it may be useful to reconsider the role of statistical analysis in the decision making process. The decision maker is concerned with choosing between two or more alternatives; the value of which remains to be established by events in the future. Statistical analysis is valuable only to the extent to which it raises the level of understanding of the problem and in so doing provides an improved basis for fixing beliefs about the future. In contrast, analyses that provide interesting expositions, but no additional understanding, are of little value. It, therefore, is essential for the analyst to be attuned to informational requirements of the decision maker if real progress is to be made.

MANAGEMENT INFORMATION SYSTEMS

A second area which could benefit from the attention of statisticians is the design of management information systems. Even a cursory examination of the recent attempts to design and implement management information systems reveals the opportunity for substantial improvement through the infusion of the concepts of experimental statistics. Many of these efforts reflect a lack of understanding of the available techniques for summarizing and analyzing data. The result of this naiveté has been inefficiency in system design and confusion regarding the purpose and value of the output of the system. For example, the operation readiness of our Hawk units throughout the world must be monitored daily by phone. Since this information is vital to decision makers at the highest levels, one would have hoped that a less cumbersome communication system could have been planned.

To provide you with more background on the problem area, it may be useful to examine briefly the origin of our current dilemma. The root of the problem can be found in our recently acquired capacity to process and transmit rapidly information. In the last thirty years technological progress has resulted in the development of three generations of computers, each of which represented a dramatic improvement over the current state-of-the-art. Equipped with the exciting abilities to process in a real time mode and to directly access data banks, the designers of these systems have moved in the direction of including everything about everything in the system.

One example of the problem is provided by the periodic Army readiness report that is prepared for the Chief-of-Staff. Included in this report, in great detail, is information on not only major items such as tanks and jeeps but also on many minor items as well. Once attention was drawn to equipment readiness at this level of specificity it became apparent that the number and status of most of the items were subject to continual change. This meant that the job of preparing a large scale report was further compounded by the fact that the information had to be updated and published frequently, if it was to be of value in its current form. A question can be asked as whether or not this is a worthwhile or even feasible effort. This same point should be raised in every management information system.

In nearly all phases of our business today one can observe information being translated into electronic impulses for transmission up to higher levels of authority. It is important to note that once data is separated from its traditional hard copy vehicle, e.g., the DA Form, it can be sorted, summarized, or transmitted at almost unbelievable speeds. It is this speed and the low per unit cost of processing information which have caused many of the current problems with management information systems.

These rapid changes in communications technology have caused some rather traumatic experiences in most large organizations. To begin with, many management theorists and most managers of today are still thinking in terms of the traditional forms of organization structure. These concepts generally involve pyramidal configurations of the different layers of authority. The problem is that these organizations reflect a certain state of information processing technology and this level of technology is rapidly becoming obsolete. There is no doubt that a certain disparity has always existed between the institutionalized organization structure and information technology; however, recent innovations have aggravated and accentuated the problem. It is useful to examine the factors that are important to this problem in order to better evaluate alternative solutions.

One important factor is the heterogeneity in the speed with which different types of information are processed through the organization. While it is not possible to rapidly analyze and summarize information on personnel strength through the organization, it is still necessary to individually monitor the progress of many R&D programs. So within the same large organization, new information processing techniques have dramatically affected the form and function of some activities while others remain essentially unchanged. This phenomenon has made the traditional concepts of a centralized and decentralized organization obsolete in that both tendencies are apparent within many phases of our business.

The increasing magnitude of the upward flow of information also serves to exacerbate the disparity between information processing technology and organizational structure. Too frequently, our concept of the informational requirements that must be transmitted up to top management reflects a lack of appreciation for the objectives of the system. Most communication that an individual has with the higher levels of the organization is through his immediate superior. Communication at this level is intimate and detailed and this is as it should be between superior and subordinate. This is not, however, the appropriate level of communication between a first line supervisor and top management. The top level manager has neither the need to know nor the capability to assimilate the large volumes of specific information, and therefore it makes little sense to send information at this level of detail up through the information system.

In addition to being illogical, this tendency has serious implications for the organization and the decision maker. If the trend continues, middle management will of necessity be relegated, in large measure, to the job of expediting the flow of information up the line of authority. More important, however, is the effect of this tendency on the performance of the decision maker. From his point of view, this tremendous flow of information provides an all encompassing yet fragmentary view of reality. While the decision maker has easy access to information regarding every significant dimension of the problem and some trivial ones as well, he may still find himself in a quandary over the nature of the situation. The reality of any situation is extremely complex when viewed in its entirety. Most of us have learned through experience in situations to suppress those aspects of reality which are superfluous to the problem at hand; however, the ability to do this effectively depends on an intimate understanding of the particular problem and environment. This point illustrates a major impetus for specialization of interest and talent but raises a serious question concerning the relationship between the top level decision maker and the information system. It is obvious that no top manager, regardless of his ability, can begin to accumulate experience comparable to the new sum of that possessed by the specialists in his organization. It should be equally obvious that the detail and format of information required by the manager is markedly different from that which is required in the lower echelons. This is, however, only half the problem.

The sorting and evaluating of information by the decision maker is further complicated by the fact that the information has been abstracted from the environment to which it is indigenous. No longer is it possible to view the situation in its totality or to make inferences from the juxtaposition of the various elements. The information is now presented in a homogenous package and there is little effort made to illustrate the relative importance of the various bits of information. This format encourages the tendency to limit the analysis to what are apparently obvious relationships in the data, and all too often, these obvious relationships depict only a superficial view of the problem. When confronted with such a situation, the decision maker is tempted to feel that his evaluation is profound when it in fact may be obvious and trivial or even worse incorrect.

The question then arises as to what alternatives are available to aid us in resolving this dilemma. One answer to the problem may be found in the imaginative and effective application of the techniques of statistical analysis. Concepts and procedures that have been used successfully for years by statisticians offer the means by which meaningful order can be restored in our information systems.

Returning to the example of the Army readiness reports, in this information system the emphasis has been placed on reporting the status of practically every item in the inventory. A moment's reflection reveals that this approach is a violation of the principle of parsimony. Why is it necessary to report data on the status of every item, when we are really only interested in those items in a particular status? It is encouraging to note that all information systems have not proceeded down the same path. The New York City Department of Public Health, for example, does not attempt to measure the health status of the city by directly estimating the proportion of the total population who are well. Instead, their attention is focused only on those who are sick. Their approach is to monitor the population of the hospitals throughout the city. Through observation of this one accessible indicator, they are able to maintain an adequate estimate of the general level of health of the community.

The principle is to replace the real variable of interest with surrogate which is more easily measured and analyzed. This has been a relatively common practice among statisticians and it should have application in the design of our information systems. In the case of the readiness report, a substantial increase in the value of the effort would be realized by reporting exceptions rather than the status of the whole system. This scheme would substantially reduce the upward flow of information and focus attention on the real variable of interest. In another phase of the operation, perhaps the status of a particular maintenance operation could be gauged more efficiently and accurately through the examination of the re-enlistment rates rather than the number of items serviced per month. The kind of changes suggested would not only reduce the upward flow of information but also place the information in a form and format that is more useful in the decision making process.

The concepts of sampling offer yet another statistical tool that appears to have application in the design of information systems. Even if modern technology can provide us with the machine capability to process information at very high speeds, this capability has a significant, positive cost. It is therefore necessary to examine alternative ways to economize in the operation. Sampling theory provides the basic notions for efficiently and economically gathering data about a particular population of interest. For example, the mean cost of procuring an item could be estimated accurately and at a mere fraction of the cost of total enumeration through the use of a self-weighting, stratified sample. It should also be remembered that in many cases, sample estimates might be even better than would usually be expected because our concern is primarily with finite populations.

A more general perspective for design of an information system may be gained from the philosophy of analysis that pervades among statisticians. While many of the designers of information systems have been content to concentrate on the preparation and reporting of data, the interest of most statisticians continues through analysis and interpretation. Efforts must be made to bring the analysis phase into the design of a system. Up to this point system designers have emphasized performance measures such as speed or cost per calculation as measures of effectiveness, but we have seen that this approach ignores the important question about system effectiveness, i.e., what is the value of information? Timeliness of information is important; however, in our effort to obtain more current data we have ignored certain other important aspects of the problem. Is it really worth anything to the organization to spend additional money to send information more quickly if much of the information in the system is already redundant or unusable? Does it make sense to publish figures in a daily report if it will require several week’s worth of observation to verify whether a change in the data is real or simply an aberration? The answer to both questions is obviously no! Both queries suggest that, in the future, major payoffs will accrue to advances in the analysis of data that can be incorporated within the system. Further analysis will take additional time; however, it should also substantially increase the informational value of reports. When examining this tradeoff it is essential to remember that most changes that take place within a large organization are gradual and occasionally painfully slow. Given this situation, it is reasonable to expect that the opportunity cost of the time lost during further analysis may be substantially less valuable than the increased understanding which would be generated.

In summary, there is a genuine need to apply the philosophy of experimental design to the design of management information and control systems. Statistical techniques can help to determine which variables should be measured and which should be ignored, as well as facilitating the analysis and forecasting of trends. Up to now, there has been little feedback between those interested in experimental design and those involved in information system design. Much of what we know in the latter area has been the result of a trial and error process, and as I am sure you are well aware, this can be a very expensive way to learn. If some of the statistical notions of sampling and analysis can be communicated to system designers, then substantial payoffs will be realized. A response in this direction now will encourage efficiency and progress. If no response is forthcoming, however, and decision making continues to escalate, a requirement for total information reporting will demand a huge organization just for purposes of processing. In many ways, the dilemma of the decision maker is analogous to that of an individual who attempts to examine the behavior of a particle suspended in liquid. The more the individual studies the particle the more confused he becomes because of the random effect of Brownian motion. The perception of both the hypothetical individual and the decision maker could be improved through the use of certain basic statistical notions.

CONCLUSION

As we have seen there are a number of opportunities to broaden the horizons of experimental design through reduced emphasis on rigor and increased attention to current problems, be they in testing or system design. The next move is up to you.