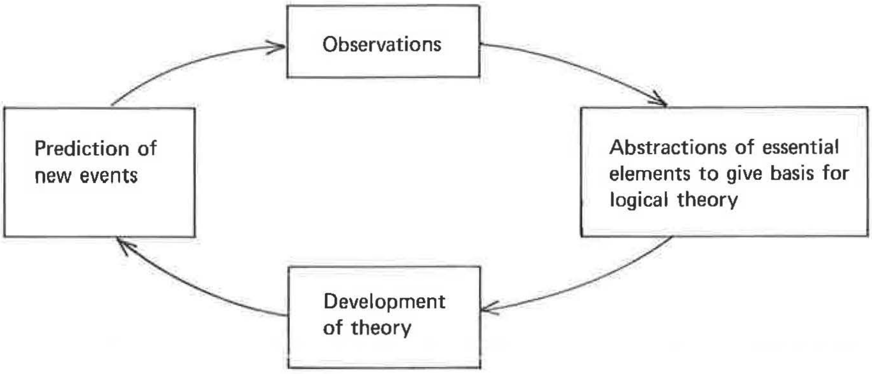
Design of Experiments

VIRGIL L. ANDERSON, Purdue University

• RESEARCH DESIGN may be considered closely allied to the scientific method. Kempthorne (1) describes the scientific method in a circular notion as follows:



The uses of statistics comes in the observations by designed experiments and analysis of data from the experiments. The design should never be considered before the analysis is contemplated.

FUNDAMENTALS OF STATISTICALLY DESIGNED EXPERIMENTS

Ingredients for Well-Designed Experiments

The first of these ingredients is to have a well-defined problem. This may take many hours, but in general an hour spent here usually saves at least ten later. In general researchers should not run one-factor-at-a-time experiments, and complete factorials are too expensive. Hence fractional factorials and modifications must be considered.

Another ingredient is to have enough replications to show "practical" significance. The notion of practical significance must be associated with cost. There is no reason to run an experiment to show statistical significance (in a probability sense) when the difference of the means is so small that it does not pay the engineer to make a change in practices. This concept can be incorporated in the experiment at the design stage if discussed thoroughly.

Two additional ingredients are to outline the analysis of the data at the design stage and to design the experiment as economically as possible.

Basic Requirements for Designs

In the past, the 2 requirements given were randomization to allow for unbiased estimates of the parameters of interest and replication to allow for an estimate of the

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

experimental error. Somehow, the engineers who have worked with me recently seem to know these, but 2 requirements are not understood.

Inference Space—The concept is the same as that of the statistician's old word "population", but engineers appreciate inference space better because it implies how far they can apply the results or what inferences they can make from the data taken from the designed experiment.

Restrictions on Randomization—In general, research workers in all fields in which I have consulted have difficulty understanding that, when a restriction on randomization is placed on the data-taking mechanism, this peculiarity must be accounted for in the analysis. It has been my experience over the past 10 years or more that almost no truly factorial experiments are run; they almost all have a tinge of "split plot" in them. I do not expect you to understand this completely here, but I will try to demonstrate it later. When canned factorial computer programs are run on your data, you will more likely misunderstand the results for your experiment than understand them if you do not understand restrictions on randomization.

Methods of Handling Extraneous Variables

Rigid Control—The inference space may be too narrow for you, but for some experiments this is desired.

Classification—This includes the treatments or factors, the blocks or restrictions on randomization, and anything else that has to do with controlling the levels of the variables to be included in the experiment.

Concomitant Information—Measuring another variable at the same time as the one of interest allows additional information with few observations.

Randomization—After rigid control, classification, and concomitant information have been used to control or measure extraneous variables to obtain unbiased estimates, randomization is used to scatter the influences of those extraneous variables missed by the first three.

Prior Information

Some statisticians do not like to use information from prior investigations, especially if randomization has been neglected. There are occasions in engineering problems, however, when prior information is necessary before an experiment can be run intelligently. If care is taken, I believe historical data can be utilized.

AN EXAMPLE

Consider the problem of deciding on the best prosthetic cardiac valve to choose from 4 types at 6 different pulse rates. An apparatus constructed by an engineer used water in the mechanical device for blood, a pump for the heart, and a certain type of tubing for blood vessels. The valves could be installed and replaced in the system with some difficulty. The pulse rates could be simulated by varying the speed of the pump quite easily.

Complete Factorial

If all 24 combinations of the valve types by pulse rates could be repeated 2 times, there would be 24 times 3 or 72 treatment combinations to run. If all of these 72 were run completely at random, as is assumed in a complete factorial, the engineer would have to disassemble his machine 71 times. The linear model to use for the analysis could be

$$Y = \mu + V + P + VP + \epsilon$$

where

Y = the variable to analyze such as back pressure or a flow variable,
μ = the overall mean,
V = the effect of the valve type,

P = the effect of the pulse rate,
VP = the interaction, and
ϵ = the error for testing the effects of V, P, and VP.

The amount of information available to investigate the valve types is associated with ϵ, which has 48 degrees of freedom. In this case the canned factorial computer program is appropriate.

Split Plot Design

Almost no engineer would take the time to disassemble the apparatus 71 times for such an experiment. He may be willing to put a valve type in the apparatus and run all 6 pulse rates at random. Then he may remove that valve type and put another one in at random. Then he may remove that valve type and put another one in at random and run all 6 pulse rates at random. He may do this for the other 2 valve types and repeat this procedure with different randomizations for 2 more times. In this case he has restricted his randomization by allowing the same valve type to receive all pulse rates before disassembling the machine. Somehow this must be accounted for in the analysis. One linear model that could show this, if certain assumptions are made, is

$$Y = \mu + V + \delta + P + VP + \epsilon'$$

In this case δ has been added in the model to account for the randomization restriction on valve types, and V is tested by δ, which has only 8 degrees of freedom. Also ϵ' has only 40 degrees of freedom, which is not much loss.

The important point to engineers is that, by making the experiment easier to run, the information on valve types, per se, may be reduced and all this must be accounted for in the analysis. In this case the canned factorial computer program is wrong and must be modified for a correct analysis.

SUMMARY

In the example, rigid control was used as much as possible by the engineer, classifications were the valve types and pulse rates, concomitant variables could be the back pressure or a flow variable, and randomization was used.

Most important, however, was that the inference space was the human beings to use these prosthetic heart valves. How immediately applicable the results would be must be ascertained by the engineers, medical doctors, and statisticians. Also, the restrictions on randomization must be accounted for if the second design (a split plot) were used. Care must be given to the analysis of such data.

REFERENCE

1. Kempthorne, O. Design and Analysis of Experiments. John Wiley and Sons, New York, 1952.