COMPLETE CLASS THEOREMS IN EXPERIMENTAL DESIGN

SYLVAIN EHRENFELD
COLUMBIA UNIVERSITY

1. Introduction

There are three broad categories into which problems of experimental design can be classified:

1) the practical problem of deciding which experiments are relevant to the problems under consideration,

2) the analysis of the particular experimental design chosen,

3) the decision as to which of the relevant experiments to perform.

Most of the work in classical design has concerned itself with the first two aspects, while the third has only recently been receiving attention. This paper deals with the third aspect.

Suppose an experimenter has available a family of random variables Y_x depending on a parameter $\theta \in \Omega \subsetneq E^{(p)}$ where $x \in A \subsetneq E^{(k)}$, with A compact and $E^{(p)}$ and $E^{(k)}$ Euclidean spaces. A choice of an experiment of size N is equivalent to choosing N points $x_1, \cdots, x_N$ lying in the set A. Performing the experiment consists in observing $Y_{x_1}, \cdots, Y_{x_N}$. If the experimenter is interested in a set of problems T, concerning the parameter θ, then the question of how to choose $x_1, \cdots, x_N$ becomes important. This is so, since the efficiency and sensitivity of the experiments with regard to the problems in the set T might be very much affected by the choice of $x_1, \cdots, x_N$.

A simple illustration is the following. Suppose Y_{x_α}, $\alpha = 1, \cdots, N$, are uncorrelated random variables with equal variance σ^2, and $E(Y_{x_\alpha}) = \beta_2 + \beta_1 x_\alpha$. The x's are assumed to be fixed constants.

It is known that the variance of the least squares estimate of β_1 is inversely proportional to $\sum (x_\alpha - \bar{x})^2$. Hence, if the values $x_1, \cdots, x_N$ can be chosen in a set $A \subsetneq E^{(1)}$, the experimenter would choose them so that $\sum_\alpha (x_\alpha - \bar{x})^2$ is as large as possible. If one were interested in β_2 as well, it is known that $x_1, \cdots, x_N$ should be chosen so that $\bar{x} = 0$. If A is the interval $-1 \leqq x \leqq 1$, and one were interested in both β_1 and β_2 then, for N even, the observations would be restricted to -1 and $+1$ with half at -1 and the other half at $+1$.

In the above the points $x_1, \cdots, x_N$ were chosen to do "well" in two problems, namely, estimating β_1 and β_2. In general the problems of interest, which we denoted by T, might include estimating certain linear relations of the form $t_1\beta_1 + t_2\beta_2$.

The experimenter can sometimes restrict himself to choosing x's in a subset of A without loss with respect to the problems in the set T. In sections 2 and 3 it will be shown how these subsets can be found.

Sponsored by the Office of Scientific Research of the Air Force, Contract AF 18 (600)-442.