OPTIMAL EXPERIMENTAL DESIGNS[1]

By Samuel Karlin[2] and William J. Studden[3]

1. Introduction. The purpose of this paper is to discuss a number of results concerning the geometric theory of the optimal design of experiments. This theory was initiated and principally developed in a series of important publications by Elfving (1952), Kiefer (1960), (1962) and Kiefer and Wolfowitz (1959), (1960). For other references and historical background we direct the reader to Kiefer (1959), (1960). This paper was motivated and inspired by numerous conversations with Kiefer and contains approximately half expository and half new material. Almost all the proofs are new and presented in a unified manner.

The theory of the optimal design of experiments fits the following structure. Let $\mathbf{f} = (f_0, f_1, \cdots, f_n)$ denote a vector of $n + 1$ linearly independent continuous functions defined on a compact space $\mathfrak{X}$. The points of $\mathfrak{X}$ are referred to as possible levels of feasible experiments. For each level $x \varepsilon \mathfrak{X}$ some experiment can be performed whose outcome is a random variable $y(x)$. It is assumed that $y(x)$ has a mean of the explicit form[4] $\mathcal{E}y(x) = \sum_{j=0}^{n} \theta_j f_j(x)$ and a common variance σ^2 independent of x (normalized for convenience $= 1$). The functions $f_0, \cdots, f_n$, called the regression functions, are assumed known to the experimenter while the parameters $\theta_0, \theta_1, \cdots, \theta_n$ are unknowns to be estimated on the basis of N *uncorrelated* observations $\{y(x_i)\}_1^N$.

An *experimental design* specifies a probability measure ξ concentrating mass $p_1, \cdots, p_r$ at the points $x_1, \cdots, x_r$ where $p_i N = n_i$, $i = 1, \cdots, r$, are integers. The associated experiment involves taking n_i observations of the random variable $y(x_i)$, $i = 1, \cdots, r$.

The problem confronting the experimenter is to choose the design possessing certain optimality properties. Statistical considerations [see Kiefer (1959)] direct an interest in the matrix $\mathbf{M}(\xi) = \|m_{ij}(\xi)\|_{i,j=0}^{n}$ ($m_{ij}(\xi) = \int_{\mathfrak{X}} f_i(x) f_j(x) \xi(dx)$) commonly called the *information matrix* of the design ξ. If the unknown parameter vector $\boldsymbol{\theta} = (\theta_0, \theta_1, \cdots, \theta_n)$ is estimated by the method of least squares thus securing a best linear unbiased estimate, say $\hat{\boldsymbol{\theta}}$, then the covariance matrix of $\hat{\boldsymbol{\theta}}$ is given by

$$\mathcal{E}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' = N^{-1}\mathbf{M}^{-1}(\xi) \tag{1.1}$$

where ξ assigns mass $p_i = n_i/N$ at the points x_i, $i = 1, \cdots, r$. If the matrix $\mathbf{M}^{-1}(\xi)$ is "small" or $\mathbf{M}(\xi)$ is "large", then roughly speaking $\hat{\boldsymbol{\theta}}$ is close to $\boldsymbol{\theta}$. Most

Received 29 September 1965.

[1] Reproduction in whole or in part permitted for any purposes of the United States Government.

[2] Work supported in part by Contracts NONR 225(28) and NSF GP 2487 at Stanford University.

[3] Work supported in part by Contract NONR 1100(26) at Purdue University.

[4] $\mathcal{E}$ denotes the expectation operator.