The Annals of Mathematical Statistics
1969, Vol. 40, No. 5, 1557–1569

ADMISSIBLE DESIGNS FOR POLYNOMIAL SPLINE REGRESSION[1]

BY W. J. STUDDEN AND D. J. VANARMAN

Purdue University

1. Introduction. Let $f = (f_0, f_1, \cdots, f_n)$ be a vector of linearly independent continuous functions on a closed interval $[a, b]$. For each x or "level" in $[a, b]$ an experiment can be performed whose outcome is a random variable $Y(x)$ with mean value $\sum_{i=0}^{n} \theta_i f_i(x)$ and variance σ^2, independent of x. The functions f_0, $f_1, \cdots, f_n$ are called the regression functions and assumed known to the experimenter while the vector of parameters $\theta = (\theta_0, \theta_1, \cdots, \theta_n)$ and σ^2 are unknown. One of the main problems in the above setup is the estimation of functions of the vector θ by means of a finite number N of uncorrelated observations $\{Y(x_1)\}_{i=1}^{N}$. Given a specific function of θ and a criterion of what a good estimate is, the design problem is one of selecting the x_i's at which to experiment. In the present paper an experimental design is a probability measure μ on $[a, b]$. The experimenter then takes his observations at the different levels proportional to the measure μ. For a more complete discussion of the above model see Kiefer (1959) or Karlin and Studden (1966a).

For estimating linear functions of θ, minimaxity problems, etc., the information matrix of μ plays an important role. For an arbitrary probability measure on $[a, b]$, the information matrix $M(\mu)$ is the matrix with elements

$$m_{ij} = m_{ij}(\mu) = \int_{[a,b]} f_i f_j \, d\mu, \qquad (i, j = 0, 1, \cdots, n).$$

For two probability measures μ and ν on $[a, b]$ we say $\nu \geqq \mu$ or $M(\nu) \geqq M(\mu)$ if the matrix $M(\nu) - M(\mu)$ is non-negative definite and unequal to the zero matrix.

DEFINITION 1. A probability measure or design μ is said to be admissible if there is no design ν such that $\nu \geqq \mu$. Otherwise μ is inadmissible.

For the case of ordinary polynomial regression where $f = (f_0, f_1, \cdots, f_n) = (1, x, \cdots, x^n)$ Kiefer (1959, page 291) has shown that μ is admissible if and only if the spectrum of μ, $S(\mu)$, has at most $n - 1$ points in the open interval (a, b). In this paper we shall generalize the above result to spline polynomial regression functions. We consider the interval $[a, b]$ and choose h fixed points or "knots" $\xi_1, \xi_2, \cdots, \xi_h$ such that $a < \xi_1 < \xi_2 < \cdots < \xi_h < b$. The type of regression function under consideration will be a polynomial of degree (at most) n on each of the $h + 1$ intervals (ξ_i, ξ_{i+1}) $i = 0, 1, \cdots, h$ ($\xi_0 = a$ and $\xi_{h+1} = b$) and will have $n - k_i - 1$ continuous derivatives at ξ_i, $i = 1, \cdots, h$. The integers k_i are assumed to satisfy $0 \leqq k_i \leqq n - 1$ so that the regression function is always at least continuous. The following lemma gives a characterization of the type of

Received 15 November 1968.

[1] This research was supported by the NSF Contract GP 8986 and by the Office of Naval Research Contract NONR 1100(26). Reproduction in whole or in part is permitted for any purposes of the United States Government.