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.

regression function we are interested in. The function x_+^j used below and throughout the paper is defined by

$$x_+^j = x^j, \quad x \geqq 0; \qquad (j = 1, 2, \cdots)$$
$$= 0, \quad x < 0,$$

LEMMA 1.1. *A function $P(x)$ on $[a, b]$ can be expressed in the form*

$$P(x) = \sum_{i=0}^{n} a_i x^i + \sum_{i=1}^{h} \sum_{j=0}^{k_i} b_{ij}(x - \xi_i)_+^{n-j} \tag{1.1}$$

if and only if

(1) *P is an ordinary polynomial of degree n in each of the intervals (ξ_i, ξ_{i+1}) $i = 0, 1, \cdots, h$ and*

(2) *P has $n - k_i - 1$ continuous derivatives at ξ_i, $i = 1, 2, \cdots, h$.*

PROOF. See Karlin and Ziegler (1966, page 518).

We shall assume that our vector of regression functions consists of the functions

$$1, x, x^2, \cdots, x^n, (x - \xi_i)_+^{n-k_i}, (x - \xi_i)_+^{n-k_i+1}, \cdots, (x - \xi_i)_+^n, \quad i = 1, 2, \cdots, h. \tag{1.2}$$

Kiefer's result mentioned above was generalized by D. VanArman (1968) who considered the cases $h = 1$ and $k_i = n - 1$ or $n - 2$ for general h. The present paper is devoted to proving the following result.

THEOREM 1.1. *Let f consist of the vector of functions in (1.2). Then a design μ is admissible if and only if the spectrum of μ, $S(\mu)$, has less than or equal to*

$$n - 1 + \sum_{j=i+1}^{i+l} [\tfrac{1}{2}(n + k_j + 1)] \tag{1.3}$$

points on the open interval (ξ_i, ξ_{i+l+1}) for $i = 0, 1, \cdots, h - l$, $l = 0, 1, \cdots, h$. (Here we let $\xi_0 = a$, $\xi_{h+1} = b$ and $[x]$ denotes the greatest integer in x.)

SPECIAL CASES. (1) If each k_j is equal to its maximal value $n - 1$, then the requirements on our regression function are that it be continuous and a polynomial of degree n on each interval (ξ_i, ξ_{i+1}) $i = 0, 1, \cdots, h$. In this case μ is admissible if and only if $S(\mu)$ has $\leqq n - 1 + ln$ points on (ξ_i, ξ_{i+l+1}). This is equivalent to $S(\mu)$ having $\leqq n - 1$ points on each (ξ_i, ξ_{i+1}), $i = 0, 1, \cdots, h$. This is a more or less direct extension of Kiefer's result.

(2) For $n = 1$ and $k_j = 0, j = 1, 2, \cdots, h$ the regression function is linear on $(\xi_j, \xi_{j+1}), j = 0, 1, \cdots, h$, and continuous at $\xi_j, j = 1, 2, \cdots, h$. By (1) above a design μ is admissible if and only if $S(\mu) \subset \{a, \xi_1, \xi_2, \cdots, \xi_h, b\}$.

(3) The first "non-trivial" case is probably $n = 2$ and $k_j = 0, j = 1, 2, \cdots, h$. Here we have a quadratic on each interval (ξ_i, ξ_{i+1}) and the regression function is required to have a continuous derivative at each ξ_j. A design μ is admissible if and only if there are at most $l + 1$ points of $S(\mu)$ in (ξ_i, ξ_{i+l+1}), $i = 0, 1, \cdots, h - l$, $l = 0, 1, \cdots, h$.

(4) Specific design problems are usually easier to analyze when the maximal number of points in any admissible design, say A, is equal to $B =$ the number of regression functions. If the regression functions are linearly independent then

$A \geqq B$. In the spline situation an admissible design may have $A = n + 1 + \sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)]$ points in $S(\mu)$. The number of regression functions is $B = n + 1 + \sum_{j=1}^{h} (k_j + 1)$. Since $[\frac{1}{2}(n + k_j + 1)] = k_j + 1 + [\frac{1}{2}(n - k_j - 1)] \geqq k_j + 1$ for all j the situation $A = B$ arises if and only if $n - k_j = 1$ or 2, $j = 1, 2, \cdots, h$. (Note that by assumption $n - k_j \geqq 1$.)

In Section 2 we give some necessary and sufficient conditions for admissibility in terms of the elements or moments in the information matrix. The proof of Theorem 1.1 is given in Section 4 after we first prove a number of preliminary lemmas which are given in Section 3.

The authors wish to thank Professors John R. Rice and Carl deBoor for a number of extremely helpful discussions concerning spline functions

2. Moment conditions for admissibility. The proof for the ordinary polynomial case uses the fact that if $f(x) = (1, x, \cdots, x^n)$ and $g(x) = (1, x, \cdots, x^{2n-1})$ then $M(\nu) \geqq M(\mu)$ if and only if $\int g(x)\, d(\nu - \mu) = 0$ and $\int x^{2n}\, d(\nu - \mu) > 0$ (all integerals in the following will be over $[a, b]$ unless specified otherwise). Thus μ is admissible if and only if the moments of μ up to order $2n - 1$ uniquely determine μ or among those measures ν whose moments up to $2n - 1$ agree with those of μ, the measure μ maximizes the $2n$th moment. The remainder of the argument then uses known moment space results which show that these latter conditions on μ are equivalent to $S(\mu)$ having at most $n - 1$ points in (a, b). The idea behind the proof of Theorem 1.1 will be to follow the above line of argument. The moment conditions for the spline polynomials are contained in Theorem 2.1 below. We let g consist of the vector of functions

$$(2.1)\quad 1, x, \cdots, x^{2n-1}, \ (x - \xi_i)_+^{n-k_i}, \cdots, (x - \xi_i)_+^{2n-1}, \quad i = 1, 2, \cdots, h.$$

THEOREM 2.1. *Let f consist of the vector of regression functions in* (1.2) *and let g be defined as above. Then $\nu \geqq \mu$* (or $M(\nu) \geqq M(\mu)$) *if and only if*

(1) $\int g(x)\, d(\nu - \mu) = 0$ *and*

(2) $0 \neq \int x^{2n}\, d(\nu - \mu) \geqq \int (x - \xi_1)_+^{2n}\, d(\nu - \mu) \geqq \cdots \geqq \int (x - \xi_h)_+^{2n}\, d(\nu - \mu) \geqq 0.$

If condition (1) *holds then condition* (2) *is equivalent to*

(2′) $\int [(x - \xi_i)_+^{2n} - (x - \xi_{i+1})_+^{2n}]\, d(\nu - \mu) \geqq 0, i = 0, 1, \cdots, h;$

and at least one of these inequalities is strict.

The proof of the above theorem will use the following two simple lemmas.

LEMMA 2.1. *If $M = (m_{ij})$ is a symmetric non-negative definite matrix and a diagonal element $m_{ii} = 0$ for some i then $m_{ij} = 0$ for all j.*

PROOF. Consider the vector with a 1 in the ith component, β in the jth component and zeros elsewhere. Then the fact that M is non-negative definite results in $2\beta m_{ij} + \beta^2 m_{jj} \geqq 0$ for all β. This readily implies that $m_{ij} = 0$.

LEMMA 2.2. *Let A be a matrix of the form*

$$A = \begin{bmatrix} A_0 & A_1 & \cdots & A_h \\ A_1 & A_1 & \cdots & A_h \\ \vdots & \vdots & & \vdots \\ A_h & A_h & \cdots & A_h \end{bmatrix}. \tag{2.2}$$

Then $A \geqq 0$ if and only if $0 \neq A_0 \geqq A_1 \geqq \cdots \geqq A_h \geqq 0$.

PROOF. We need only notice that

$$\textstyle\sum_{i,j} x_i x_j a_{ij} = \sum_{i=0}^{h} (A_i - A_{i+1})(x_1 + \cdots + x_i)^2$$

where $A_{h+1} = 0$.

PROOF OF THEOREM 2.1. The vector f of regression functions specified in (1.2) will be written with the powers $1, x, \cdots, x^n$ in the first $n + 1$ components, then those involving ξ_1, etc. We let $M = M(\nu) - M(\mu)$ and assume that $M \geqq 0$. The proof that conditions (1) and (2) hold consists of a repeated application of Lemma 2.1 and a single application of Lemma 2.2. Since ν and μ are both probability measures the first row (and column) of M has zero elements, i.e.

(a) $\int x^i \, d(\nu - \mu) = 0, i = 1, 2, \cdots, n.$

(b) $\int (x - \xi_p)_+^j \, d(\nu - \mu) = 0, j = n - k_p, \cdots, n; p = 1, \cdots, h.$

From (a) with $i = 2$, the 2nd row is also zero. Finally we obtain

$$\int x^i \, d(\nu - \mu) = 0, \quad i = 0, 1, \cdots, 2n - 1;$$

$$\int x^i (x - \xi_p)_+^j \, d(\nu - \mu) = 0, \quad i = 0, 1, \cdots, n - 1; \quad j = n - k_p, \cdots, n; \quad p = 1, 2, \cdots, h.$$

Now for $r \leqq n - 1$ and any $p = 1, 2, \cdots, h$,

$$\begin{aligned} &\int (x - \xi_p)_+^{n+r} \, d(\nu - \mu) \\ &\quad = \int (x - \xi_p)_+^r (x - \xi_p)_+^n \, d(\nu - \mu) = \textstyle\sum_{i=0}^{r} a_i \int x^i (x - \xi_p)_+^n \, d(\nu - \mu) = 0. \end{aligned}$$

Therefore $\int g(x) \, d(\nu - \mu) = 0$, i.e. condition (1) holds. Now observe that M has all diagonal elements equal to zero except the elements

$$\int x^{2n} \, d(\nu - \mu) \quad \text{and} \quad \int (x - \xi_p)_+^{2n} \, d(\nu - \mu), \qquad p = 1, \cdots, h,$$

and M has zero elements except for the corresponding rows and columns. We now show that the resulting submatrix has the form (2.2). The element $A_0 = \int x^{2n} \, d(\nu - \mu)$. The element A_1 in the first row is

$$\int x^n (x - \xi_1)_+^n \, d(\nu - \mu) = \int (x - \xi_1)_+^{2n} \, d(\nu - \mu).$$

Similarly the element, say, in the 2nd row and 3rd column is

$$\int (x - \xi_1)_+^n (x - \xi_2)_+^n \, d(\nu - \mu) = \int (x - \xi_2)_+^{2n} \, d(\nu - \mu).$$

Applying Lemma 2.2 we find that condition (2) holds.

Now assume that conditions (1) and (2) hold. Condition (1) reduces M as above so that the nonzero part of M has the form (2.2). Condition (2) and Lemma 2.2 then imply that $M \geqq 0$.

For a design to be admissible a trivial necessary condition is that given any subset I of $[a, b]$ the design μ, normalized on I, is admissible on I relative to the same regression vector. An explicit definition and lemma in this regard is given below.

DEFINITION 2.1. A design μ is said to be admissible on $[\xi_i, \xi_{i+l+1}]$ if μ (normalized on this interval) is admissible there relative to the functions

$$(2.3)\quad 1, x, \cdots, x^n, (x - \xi_j)^{n-k_j}, \cdots, (x - \xi_j)_+^n, \qquad j = i + 1, \cdots, i + l.$$

A design μ is said to be subadmissible (h) if it is admissible on every subinterval $[\xi_i, \xi_{i+l+1}]$, $i = 0, 1, \cdots, h - l$, $l = 0, 1, \cdots, h - 1$.

LEMMA 2.3. *If μ is admissible for h knots then μ is subadmissible.*

PROOF. By the remarks preceding Definition 2.1 the measure μ must be admissible on $[\xi_i, \xi_{i+l+1}]$ relative to the functions (1.2) restricted to this interval. Admissibility is independent of the basis used for the linear space spanned by the functions. By Lemma 1.1 the functions (2.3) are a basis for the linear space spanned by the functions (1.2) restricted to $[\xi_i, \xi_{i+l+1}]$.

3. Some preliminary lemmas. In this section we present a number of lemmas which will be used in the proof of Theorem 1.1. These lemmas deal with the construction of some spline polynomials having specified zeros.

First we paraphrase an important result of Karlin and Ziegler (1966, pages 519–522). (See also Karlin (1968).) Let $\varphi_s(t_i, u_j) = (t_i - u_j)_+^s$, $s = 1, 2, \cdots$ and let $t_i, u_j, i, j = 1, 2, \cdots, r$ satisfy the following conditions:

(1) $c \leqq t_i \leqq \cdots \leqq t_r \leqq d; \qquad c \leqq u_1 \leqq \cdots \leqq u_r \leqq d.$

(2) $\alpha + \beta \leqq s + 2$ $(s \geqq 1)$ whenever $\alpha (\geqq 1)$ of the u_j's coincide, say equal to g, and $\beta (\geqq 1)$ of the t_i's agree with the same point g.

(3) No more than $s + 1$ consecutive t_i's (or u_j's) coincide.

Let $M_s(t, u)$ be defined as follows: if $t_1 < t_2 < \cdots < t_r$ and $u_1 < u_2 < \cdots < u_r$, $M_s(t, u)$ is the matrix $\| \varphi_s(t_i, u_j) \|_{i,j=1}^r$. If $u_{j_0-1} < u_{j_0} = u_{j_0+1} = \cdots = u_{j_0+h-1} < u_{j_0+h}$ we replace the $(j_0 + i)$th column vector, $1 \leqq i \leqq h - 1$, of $\|\varphi_s(t_i, u_j)\|_{i,j-1}^r$ by $(d^i\varphi_s(t_v, u)/du^i)_{u=u_{j_0}}$, $v = 1, \cdots, r$. A similar adjustment is used on the rows of the matrix when t_i values coincide, any sth derivative being taken from the right. We let $D_s(t, u)$ be the determinant of $M_s(t, u)$. The result of Karlin and Ziegler is that under conditions (1), (2) and (3), $D_s(t, u) \geqq 0$ always and

$$(3.1)\qquad D_s(t, u) > 0 \Leftrightarrow t_{i-s-1} < u_i < t_i, \qquad i = 1, 2, \cdots, r$$

where for $i \leqq s + 1$ only the right hand inequality is relevant. For any vector of functions $f(x) = (f_1(x), \cdots, f_h(x))$ and vector of constants $t = (t_1, \cdots, t_h)$ where $t_1 \leqq t_2 \leqq \cdots \leqq t_h$ we let $M(t, f)$ be the matrix with the vector $f(t_i)$ in the

ith row. If t_i values coincide then the successive rows are replaced by derivatives as in the definition of $M_s(t, u)$.

LEMMA 3.1. *Let f denote the vector of functions*

$$1, x, \cdots, x^s, (x - \xi_i)_+^{s-\lambda_i}, \cdots, (x - \xi_i)_+^s \quad i = 1, \cdots, h; \tag{3.2}$$

where $0 \leqq \lambda_i \leqq s - 1; i = 1, \cdots, h$.
Let $t = (t_1, \cdots, t_r)$ *where* $r = s + 1 + h + \sum_{j=1}^h \lambda_j$, *no more than* $(s - \lambda_i + 1)\, t_j$ *values are* ξ_i, *and no more than* $(s + 1)\, t_j$ *values coincide. Then* $M(t, f)$ *is non-singular if and only if*

$$t_{\gamma_i} < \xi_i < t_{s+2+\gamma_{i-1}} \qquad i = 1, 2, \cdots, h \tag{3.3}$$

where $\gamma_i = \sum_{j=1}^i (\lambda_j + 1), i = 1, 2, \cdots, h, \gamma_0 = 0$.

PROOF. The result is an application of the Karlin-Ziegler result with $c = a - \epsilon$, $d = b$, where for the u_j values we choose $s + 1$ equal to c and $\lambda_i + 1$ equal to $\xi_i, i = 1, \cdots, h$.

It is then seen that the matrices $M_s(t, u)$ and $M(t, f)$ are non-singular together. The inequalities (3.3) are equivalent to those in (3.1).

The proof of our main theorem will require a somewhat delicate analysis of the zeros of polynomials constructed using Lemma 3.1. This is due mainly to the fact that spline polynomials are not infinitely differentiable and non-trivial spline polynomials may vanish identically on intervals between knots. All systems of functions we shall use will be linearly independent so that a linear combination of these functions will be trivial or identically zero on $(-\infty, \infty)$ if and only if all the coefficients vanish.

We shall use the following conventions when counting the zeros of a spline polynomial $P(x)$. (See Karlin and Schumaker (1967).)

(a) no zeros are counted on any open interval (ξ_i, ξ_j) if $P(x) \equiv 0$ there.

(b) the multiplicity of a zero $z \neq \xi_i, i = 1, 2, \cdots, h$ is counted in the usual manner, i.e., z is a zero of order r if

$$P^{(j)}(z) = 0, \qquad j = 0, 1, \cdots, r - 1, \qquad P^{(r)}(z) \neq 0.$$

(c) if $P(x) \equiv 0$ on (ξ_{i-1}, ξ_i) and $\not\equiv 0$ on (ξ_i, ξ_{i+1}) the zero at ξ_i is counted as in (b) using right hand derivatives. Similarly we use left hand derivatives for $P(x) \not\equiv 0$ on (ξ_{i-1}, ξ_i) and $\equiv 0$ on (ξ_i, ξ_{i+1}).

(d) If $P(x) \not\equiv 0$ on (ξ_{i-1}, ξ_i) or (ξ_i, ξ_{i+1}) and

$$P^{(j)}(\xi_i-) = P^{(j)}(\xi_i+) = 0, \quad j = 0, 1, \cdots, r - 1 \tag{3.4}$$

$$A = P^{(r)}(\xi_i-) \neq P^{(r)}(\xi_i+) = B$$

then ξ_i is a zero of order

(i) r if $AB > 0$;
(ii) $r + 1$ if $AB < 0$;
(iii) $r + 1$ if $AB = 0$ and $B - A > 0$;
$r + 2$ if $AB = 0$ and $B - A < 0$.

It is easily seen that a zero of order r of $P(x)$ is a zero of order $r - 1$ of P'. We let $Z(P)$ denote the number of zeros of P according to the above conventions.

LEMMA 3.2. (i) *A non-trivial polynomial P in the functions*

$$(3.5) \qquad 1, x, \cdots, x^s, (x - \xi_j)_+^{p_j}, \cdots, (x - \xi_j)_+^s, \quad j = 1, 2, \cdots, h;$$

where $1 \leqq p_j \leqq s$, *has* $Z(P) \leqq s + \sum_{j=1}^h (s - p_j + 1)$.

(ii) *For any fixed* $i = 0, 1, \cdots, h$ *a non-trivial polynomial P in the functions in* (3.5) *and* $f_i(x) = (x - \xi_i)_+^{s+1} - (x - \xi_{i+1})_+^{s+1}$ *satisfies* $Z(P) \leqq s + 1 + \sum_{j=1}^h (s - p_j + 1)$ *zeros.*

(*Note that in each case the maximal number of zeros is one less than number of functions used.*)

PROOF. (i) The proof will proceed using an induction on h and p_j and an application of Rolle's Theorem. For $h = 0$, when no spline parts are present, we are in the ordinary polynomial case and the result is immediate. For $h = 1$ we consider first the case $p_1 = 1$, where only continuity of P is required at ξ_1. If $P(x) \equiv 0$ on (a, ξ_1) or (ξ_1, b) then $Z(P) \leqq s$. If $P(x) \not\equiv 0$ on (a, ξ_1) or (ξ_1, b) and $P(\xi_1) \neq 0$ then $Z(P) \leqq 2s$. Moreover if (3.4) holds and $AB \neq 0$ then P has at most $r + 1$ zeros at ξ_1, and at most $s - r$ on each side of ξ_1 so $Z(P) \leqq 2(s - r) + r - 1 = 2s + 1 - r \leqq 2s$. If $AB = 0$ then P has at most $r + 2$ zeros at ξ_1 and at most $(s - r) + (s - r - 1)$ zeros not equal to ξ_1 so that again $Z(P) \leqq 2s$.

Still letting $h = 1$ we now assume the result true for any s and $p_1 = 1, 2, \cdots, i - 1$ and consider the case $p_1 = i$. If $Z(P) > 2s - i + 1$ then by Rolle's Theorem $Z(P') > 2s - i$. To deduce this, care must be taken with the intervals on which P or P' vanishes. For example if $P \not\equiv 0$ on (a, ξ_1) or (ξ_1, b) but $P' \equiv 0$ on (a, ξ_1) or (ξ_1, b) then P is constant on the corresponding interval. The induction hypothesis now furnishes a contradiction since $Z(P') \leqq 2s - i$.

We now apply an induction on h and $p_{j_0} = \min p_j$. We first observe (see proof of Lemma 2.3) that a basis for the linear space spanned by the functions (3.5) restricted to (ξ_i, ξ_j) is the system (3.5) omitting the spline parts involving $\xi_1, \cdots, \xi_i, \xi_j, \cdots, \xi_h$. If $p_{j_0} = 1$ we proceed as in the case $h = 1$, $p_1 = 1$ using the induction on h. We then suppose the result true for all s, h and $p_{j_0} = 1, 2, \cdots, i - 1$ and deduce the result for $p_{j_0} = i$ again using Rolle's Theorem.

(ii) The arguments used here are similar to those used in part (i). The induction steps are carried out in the same order and will be omitted.

LEMMA 3.3. *Consider the functions* (3.5) *where* $s = 2n$ *and for* $i = 1, \cdots, h$,

$$(3.6) \qquad \begin{aligned} p_i &= n - k_i, && n - k_i \text{ odd}; \\ &= n - k_i + 1, && n - k_i \text{ even}. \end{aligned}$$

Let $\lambda_i = 2n - p_i$, *assume* $p_i \geqq 3$ *and consider a set of t values* $a = t_1 < t_2 \leqq t_3 \leqq \cdots \leqq t_{m-1} < t_m = b$ *were* $m = 2n + \sum_{j=1}^h (\lambda_j + 1)$ *and at most two of the* t_i *values are equal to any given point. If*

$$(3.7) \qquad t_{\gamma_i} < \xi_i < t_{2n+1+\gamma_{i-1}}, \qquad i = 1, 2, \cdots, h,$$

where $\gamma_i = \sum_{j=1}^{i} (\lambda_j + 1)$, $i = 1, 2, \cdots, h$, $\gamma_0 = 0$, *then there exists a polynomial* $P(x)$ *in the functions* (3.5) *such that*

(i) $P(t_i) = 0$, $i = 1, 2, \cdots, m$ *with double zeros at* t_i *if* $t_i = t_{i+1}$

(ii) $P(x) \not\equiv 0$ *on any subinterval*

(iii) $P(b + 1) = -1$

(iv) *if* b_j, $j = 0, 1, \cdots, h$ *denotes the coefficient of* x^{2n} *on* (ξ_j, ξ_{j+1}) *then* $b_j \leqq 0$. (*Note that since* (ii) *holds all of the* t_i *values may be counted as zeros of* P, *i.e.* $Z(P) = m$.)

PROOF. The polynomial is constructed using Lemma 3.1. The m t_j values and $t_{m+1} = b + 1$ are used so that the r in Lemma 3.1 is $m + 1$. We set up the system of equations for P by requiring that $P(t_i) = 0, i = 1, \cdots, m$ (appropriate derivatives if t_i values coincide) and $P(b + 1) = -1$. By Lemma 3.1 the resulting system of equations has a non-vanishing determinant. Therefore the polynomial P exists and conditions (i) and (iii) are satisfied. In order to prove (ii) it suffices to show that P cannot vanish at a value t_0 distinct from the t_i, $i = 1, \cdots, m$. If P did vanish at some other point we consider the set $t_0, t_1, \cdots, t_m$ in non-decreasing order. If we renumber this set the new subscripts can be increased by at most one. Thus for the new system the inequalities (3.3) hold since we have shifted the value $s + 2$ in (3.3) to $2n + 1 = s + 1$ in (3.7). Therefore the polynomial P is identically zero which is a contradiction.

Now consider part (iv). Let $p = p_{i_0} = \min p_i$ and consider the derivative polynomial $P^{(p-1)}(x)$ whose highest coefficient is $2n - (p - 1)$. $P(x)$ has $2n + \sum_{i=1}^{h} (\lambda_i + 1)$ zeros so that $P^{(p-1)}(x)$ has at least $2n + \sum_{i=1}^{h} (\lambda_i + 1) - (p - 1)$ distinct zeros. Since $P(x)$ does not vanish on any subinterval we may assume that none of these zeros are counted in any open interval on which $P^{(p-1)}(x)$ vanishes identically. Applying Lemma 3.2 we see that $P^{(p-1)}(x)$ has

$$(3.8) \qquad \leqq 2n - p_{i_0} + 1 + \sum_{j-1}^{i_0-1} (\lambda_j + 1) \text{ zeros on } [a, \xi_{i_0}], \text{ and}$$

$$(3.9) \qquad \leqq 2n - p_{i_0} + 1 + \sum_{j=i_0+1}^{h} (\lambda_j + 1) \text{ zeros on } [\xi_{i_0}, b].$$

Adding these two numbers we find their sum to be exactly $2n + \sum_{j=1}^{h} (\lambda_j + 1) - (p_{i_0} - 1)$ so that equality must occur in (3.8) and (3.9). Moreover these are all distinct zeros so that if ξ_{i_0} is a zero it can be counted in only one of the intervals $[a, \xi_{i_0}]$ and $[\xi_{i_0}, b]$. In this case (3.8) and (3.9) may be written using $[a, \xi_{i_0})$ and $(\xi_{i_0}, b]$.

If $b_h > 0$, then $P(x) \to \infty$ as $x \to +\infty$. Since $P(b + 1) = -1$, the polynomial P then has a zero above $b + 1$. However P has a maximal number of zeros on $[a, b]$. Therefore $b_h \leqq 0$. Since the number m is even, $P(a - 1) < 0$, so that we also have $b_0 \leqq 0$. Using the maximal number of zeros of each derivative we see that $P^{(p-1)}(x) < 0$ to the right of its largest zero on (a, b). Since the number of zeros of $P^{(p-1)}(x)$ on (ξ_{i_0}, b) given by (3.9) is even we conclude that $P^{(p-1)}(\xi_{i_0}) < 0$. We extend $P^{(p-1)}(x)$ to the left of ξ_{i_0} using its expression on (ξ_{i_0}, ξ_{i_0+1}). Since this new polynomial already has a maximal number of zeros on (ξ_{i_0}, b) it follows that $b_{i_0} \leqq 0$. Similarly $b_{i_0-1} \leqq 0$. We now consider the new polynomial and differentiate up to the value $\min_{i_0<j\leqq h} p_j$, thus deducing, as

above, that additional b_j values for $i_0 < j \leqq h$ are $\leqq 0$. Continuing in this manner on both sides of ξ_{i_0} we may conclude that $b_j \leqq 0, 0 \leqq j \leqq h$.

LEMMA 3.4. *Let $S(\mu)$ consist of $n + \sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)]$ points in the open interval (a, b) and suppose $S(\mu)$ has less than or equal to*

$$n - 1 + \sum_{j=i+1}^{i+l} [\tfrac{1}{2}(n + k_j + 1)]$$

points on the open interval (ξ_i, ξ_{i+l+1}) for $i = 0, 1, \cdots, h - l, l = 0, 1, \cdots, h - 1$. Then there exists a set of polynomials $\{P_i(x)\}_{i=0}^{h}$ (one for each interval) where P_i is a polynomial in the functions

$$1, x, \cdots, x^{2n-1}, (x, \xi_j)_+^{q_j}, \cdots, (x - \xi_j)^{2n-1}, \quad j = 1, 2, \cdots, h; \tag{3.10}$$

where

$$\begin{aligned} q_j &= n - k_j - 1, \quad n - k_j \ \textit{odd}; \qquad j = 1, 2, \cdots, h \\ &= n - k_j, \quad n - k_j \ \textit{even}; \end{aligned}$$

and the function

$$f_i(x) = (x - \xi_i)_+^{2n} - (x - \xi_{i+1})_+^{2n}$$

such that

(1) *the coefficient of f_i is one,*

(2) *$P_i(x) = 0$ for $x \varepsilon S(\mu)$,*

(3) *$P_i(x) \geqq 0$ for all x,*

(4) *$P_i(x) > 0$ for $x \varepsilon [\xi_i, \xi_{i+1}], x \not\varepsilon S(\mu)$.*

REMARK. It can readily be seen that the conditions on $S(\mu)$ are incompatible if $k_j = n - 1$ for some j. In this case $q_j \geqq 2$ for all j. Further, if $k_j = n - 2$ for some j then $\xi_j \varepsilon S(\mu)$. The polynomial we construct is actually unique so that if $k_j = n - 2$ then $P_i(x) = 0$ for $x > \xi_j$ if $j \geqq i + 1$ and $P_i(x) = 0$ for $x < \xi_j$ if $j \leqq i$. For example if $k_j = n - 2$ for all j then $S(\mu)$ must have $n - 1$ points in (a, ξ_1) and (ξ_h, b), $n - 2$ points in each $(\xi_j, \xi_{j+1}), j = 1, \cdots, h - 1$ and one point at each ξ_j. The polynomial $P_i(x)$ in this case vanishes on $[a, b] \cap (\xi_i, \xi_{i+1})^c$.

PROOF. We consider a sequence $t_1, \cdots, t_m$ in non-decreasing order, where $m = 2n + 2\sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)]$ and $t_1 = t_2 =$ 1st point in $S(\mu)$, $t_3 = t_4 =$ 2nd point, etc. We construct the polynomial by taking a linear combination of the functions (3.10) and equate it and its derivative to $-f_i$ at the points in $S(\mu)$. We let $\lambda_j = 2n - 1 - q_j, j = 1, 2, \cdots, h$. The conditions of the lemma guarantee that (3.3) holds with $s = 2n - 1$. Therefore the determinant of the resulting system of equations is non-zero by Lemma 3.1. Thus a polynomial P_i in the functions (3.10) and f_i exists satisfying (1) and (2). Moreover it is unique and has at least a double zero at each point in $S(\mu)$.

Suppose that $P_i(x) \equiv 0$ for $x \varepsilon [\xi_j, \xi_{j+1}]$ where $h \geqq j \geqq i + 1$. If we set

$$\begin{aligned} Q(x) &= P_i(x), \quad x \leqq \xi_j; \\ &= 0, \quad x > \xi_j; \end{aligned}$$

then Q satisfies (1) and (2) and is a linear combination of (3.10) and f_i. Since P_i is unique $P_i(x) = 0$ for all $x > \xi_j$. A similar result holds for $j + 1 \leqq i$. Thus there is a maximal interval $[\xi_\alpha, \xi_\beta] \subset [a, b]$, which contains $[\xi_i, \xi_{i+1}]$, such that $P_i \not\equiv 0$ on any subinterval.

We now proceed as in the proof of Lemma 3.3 confining ourselves to the subinterval $[\xi_\alpha, \xi_\beta]$. Observe that if $\alpha \geqq 1 (\beta \leqq h)$ then $P_i(x)$ has a zero of order $q_\alpha(q_\beta)$ at $\xi_\alpha(\xi_\beta)$.

We shall suppose that $\alpha = 0$ and $\beta \leqq h$; the other cases may be treated in a similar manner. Since (3.3) holds P_i has at least $\sum_{j=1}^{\beta} (\lambda_j + 1)$ zeros on (a, ξ_β) and q_β zeros at ξ_β. Moreover $\lambda_\beta + q_\beta = 2n$ so that P_i has at least $2n + 1 + \sum_{j=1}^{\beta-1} (\lambda_j + 1)$ zeros which is the maximal number allowed by Lemma 3.2. Therefore $P_i(x) \neq 0$ on $[a, \xi_\beta)$ for $x \neq t_j$. Now if $\xi_\beta \not\in S(\mu)$ then there are at most $A = 2n - 2 + \sum_{j=\beta+1}^{h} (\lambda_j + 1)\ t_j$ values in (ξ_β, b) and hence P_i has at least $2n + \sum_{j=1}^{h} (\lambda_j + 1) - A = 2 + \sum_{j=1}^{\beta} (\lambda_j + 1)$ zeros in (a, ξ_β). In this case P_i has $2n + 3 + \sum_{j=1}^{\beta-1} (\lambda_j + 1)$ zeros in $(a, \xi_\beta]$ contradicting Lemma 3.2. Thus $P_i(x) \neq 0$ on the closed interval $[a, \xi_\beta]$ provided $x \neq t_j$. Therefore condition (4) is true provided we can show that (3) holds.

We modify the polynomial P_i by omitting the spline parts $(x - \xi_j)_+^k$ for $j \geqq \beta$. Suppose $\beta = i + 1$. Since q_β is even and the coefficient of f_i is one the polynomial P_i must be non-negative; otherwise the modified P_i would have an additional zero on $[\xi_\beta, \infty)$. Therefore assume $\beta > i + 1$ and let $q = q_{j_0} = \min_{1 \leqq j < \beta} q_j$. As in Lemma 3.3 we may deduce that the number of zeros of $P_i^{(q-1)}(x)$ on each of the intervals (a, ξ_{j_0}) and (ξ_{j_0}, ξ_β) is the maximal number allowed by Lemma 3.2. Continuing up to the next smallest value of $q_j - 1$, etc., we finally conclude that $P_i^{(r-1)}(x)(r = q_{\beta-1})$ has a maximal number of zeros on $(\xi_{\beta-1}, \xi_\beta]$. Therefore the coefficient of x^{2n-1} in $P_i(x)$ on $(\xi_{\beta-1}, \xi_\beta)$ cannot be zero. If $P_i(x) \leqq 0$ this coefficient must be <0 since, as above, q_β is even and the modified P_i already has a maximal number of zeros on $(a, \xi_\beta]$. Now consider $q = q_{j_0} = \min_{1 \leqq j < \beta} q_j$ again. If $(\xi_i, \xi_{i+1}) \subset (a, \xi_{j_0})$ then $P_i^{(q-1)}$ has an even number of zeros on (ξ_{j_0}, b) and is <0 for $x > \xi_\beta$ if $P_i(x) \leqq 0$. In this case $P_i^{(q-1)}(\xi_{j_0}) < 0$. If $(\xi_i, \xi_{i+1}) \subset (\xi_{j_0}, b)$ then $P_i^{(q-1)}(x) < 0$ for $x > \xi_\beta$. We take $P_i^{(q-1)}$ on whichever subinterval contains (ξ_i, ξ_{i+1}). In either case $P_i^{(q-1)}(x) < 0$ to the right of the endpoint. Now if $j_0 = i + 1$ we obtain a contradiction since $P_i^{(q-1)}$ extended to the right of ξ_{i+1} has a highest coefficient which is even and positive. If $j_0 \neq i + 1$ we continue differentiating on the subinterval containing (ξ_i, ξ_{i+1}). Eventually we arrive at a contradiction as above.

Lemma 3.5. *Let $S(\mu)$ satisfy the hypothesis of Lemma* 3.4 *with $k_j \leqq n - 2$ (or $q_j \geqq 2$) and let g denote the vector* of *functions* (3.10). *Then there exists a measure ν such that $\int g\, d(\nu - \mu) = 0$ and $S(\nu) \not\subset S(\mu)$.*

Proof. The proof proceeds along the lines of the proof in Karlin and Studden (1966b, pages 138–139) and we shall be brief. The system of functions (3.10) is a WT-system. A perturbation of the vector with a Gaussian kernel produces a T-system. The measure μ is an "upper principal representation" of the corresponding moment vector and the measure ν is a limit of measures with mass at the endpoints a and b and $n - 1 + \sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)]$ points of (a, b). Thus $\nu \neq \mu$. Moreover $S(\nu) \subset S(\mu)$ is readily seen to imply that $\nu = \mu$ since $\int g\, d(\nu - \mu) = 0$.

4. Proof of Theorem 1.1. The proof of Theorem 1.1 is a combination of the following two lemmas.

LEMMA 4.1. *A design* μ *is admissible if*

$$(4.1)\quad S(\mu) \quad \textit{has} \quad \leqq n - 1 + \textstyle\sum_{j=i+1}^{i+l} [\frac{1}{2}(n + k_j + 1)] \quad \textit{points on}$$

$$(\xi_i, \xi_{i+l+1}) \quad \textit{for} \quad i = 0, 1, \cdots, h - l; \qquad l = 0, 1, \cdots, h.$$

PROOF. Since a subspectrum of an admissible spectrum is admissible it suffices to consider the case where $S(\mu)$ satisfies (4.1) for $l = 0, 1, \cdots, h - 1$, equality holds for $l = h$ and both of the endpoints are in $S(\mu)$.

We consider a sequence of points

$$t_1, t_2, \cdots, t_m, \qquad (m = 2n + 2 \textstyle\sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)])$$

where $t_1 = a$, $t_2 = t_3 =$ 1st point of $S(\mu)$ in (a, b), $\cdots$, $t_m = b$. Let $\lambda_j + 1 = 2[\frac{1}{2}(n + k_j + 1)]$, $j = 1, 2, \cdots, h$. By condition (4.1) there are $\leqq 2n - 2 + \sum_{j=1}^{i-1} (\lambda_j + 1)$ t values in (a, ξ_i) and $\leqq 2n - 2 + \sum_{j=i+1}^{h} (\lambda_j + 1)$ t values in (ξ_i, b). Therefore if $\gamma_i = \sum_{j=1}^{i} (\lambda_j + 1)$ then

$$(4.2)\qquad \xi_i \leqq t_{2n+\gamma_{i-1}} = t_{2n+1+\gamma_{i-1}}$$

and

$$(4.3)\qquad t_{\gamma_i} = t_{1+\gamma_i} \leqq \xi_i .$$

If equality occurs in (4.2) we shift the odd numbered point $t_{2n+1+\gamma_{i-1}}$ so that it is greater than ξ_i. Similarly we shift the even numbered point so that $t_{\gamma_i} < \xi_i$ if equality occurs in (4.3). We shall assume for the moment that $k_i \leqq n - 2$. In this case $S(\mu)$ is contained in the new sequence of points since in order to shift a double t value both right and left we need $\gamma_i = 2n + \gamma_{i-1}$ or $\lambda_i + 1 = 2n$. However $\lambda_i + 1 \leqq n + k_i + 1 \leqq 2n - 1$ (provided $k_i \leqq n - 2$). The new t sequence satisfies

$$t_{\gamma_i} < \xi_i < t_{2n+1+\gamma_{i-1}}, \qquad i = 1, 2, \cdots, h.$$

Then, by Lemma 3.3, there exists a polynomial $P(x)$ in the functions

$$1, x, \cdots, x^{2n}, (x - \xi_i)_+^{p_i}, \cdots, (x - \xi_i)^{2n}, \qquad i = 1, \cdots, h;$$

where

$$p_i = n - k_i, \qquad n - k_i \text{ odd};$$
$$= n - k_i + 1, \qquad n - k_i \text{ even};$$

such that (i) $P(x)$ vanishes only on the modified set of t_j values.

(ii) $P(x) \geqq 0$ on (a, b) except between modified t values.

(iii) the coefficient of x^{2n} on each (ξ_i, ξ_{i+1}) is $\leqq 0$.

If $k_i = n - 1$ for some i the conditions (4.1) imply that $S(\mu)$ has $n - 1 + \sum_{j=1}^{i-1} [\frac{1}{2}(n + k_j + 1)]$ points in (a, ξ_i) and $n - 1 + \sum_{j=i+1}^{h} [\frac{1}{2}(n + k_j + 1)]$ points in (ξ_i, b) and a point at ξ_i. We then use the above result to construct the polynomial P on the segments where $k_j \geqq n - 2$ separately. The "combined" polynomial satisfies (i), (ii), and (iii) above.

Now condition (iii) allows us to write

$$(4.4)\quad P(x) = \sum_{j=0}^{h} a_j(x - \xi_j)_+^{2n} + R(x)$$
$$= \sum_{j=0}^{h} (a_0 + \cdots + a_j)((x - \xi_i)_+^{2n} - (x - \xi_{i+1})_+^{2n}) + R(x)$$

where $a_0 + \cdots + a_j \leqq 0$. The part involving $R(x)$ involves powers $\leqq 2n - 1$ so that if $\nu \geqq \mu$ then by Theorem 2.1, part 1, $\int R\, d(\nu - \mu) = 0$. Also $P(x)$ vanishes on $S(\mu)$ so $\int P\, d\mu = 0$. Therefore if $\nu \geqq \mu$ part 2 of Theorem 2.1 and the expression (4.4) implies that $\int P\, d\nu \leqq 0$. Now we may assume that the t_j values were modified so that $P(x) \geqq 0$ on $S(\nu)$. Since $P(x)$ vanishes only on the modified t_j set this implies that $S(\nu) \subset$ modified t_j set, which, in turn, implies that $\nu = \mu$ by Lemma 3.1. Thus μ is admissible.

LEMMA 4.2. *A design μ such that $S(\mu)$ has $\geqq n + \sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)]$ points in (a, b) is inadmissible.*

PROOF. We may assume that $S(\mu)$ consists solely of exactly

$$n + \sum_{j=1}^{h} [\tfrac{1}{2}(n + k_j + 1)]$$

points in (a, b). We proceed by induction on h. Note the result is true for $h = 0$. We assume the result true for $0, 1, \cdots, h - 1$. Then we may also assume that $S(\mu)$ satisfies (4.1) for $l = 0, 1, \cdots, h - 1$; otherwise μ is not subadmissible and hence not admissible by Lemma 2.3. Let $g(x)$ consist of the vector with component functions

$$1, \text{x}, \cdots, x^{2n-1}, \qquad (x - \xi_i)_+^{p_i}, \cdots, (x - \xi_i)_+^{2n-1}, \qquad i = 1, 2, \cdots, h;$$

where

$$p_i = n - k_i, \qquad n - k_i \text{ even};$$
$$= n - k_i - 1, \qquad n - k_i \text{ odd}.$$

If $k_i = n - 1$ for some i, the conditions (4.1) for $l = 0, \cdots, h - 1$ and the fact that $S(\mu)$ has $n + \sum_{j=1}^{h} [\frac{1}{2}(n + k_j + 1)]$ points in (a, b), imply that $\xi_i \varepsilon S(\mu)$ and μ is inadmissible on (a, ξ_i) or (ξ_i, b). Thus we assume $k_i \leqq n - 2$ so $p_i \geqq 2$. Consider the ν of Lemma 3.5 and the polynomials $\{P_i(x)\}_{i=0}^{h}$ of Lemma 3.4. Then $\int P_i\, d(\nu - \mu) = \int f_i(x)\, d(\nu - \mu) \geqq 0, i = 0, \cdots, h$, and $\int f_i(x)\, d(\nu - \mu) > 0$ for some i. Also $\int g\, d(\nu - \mu) = 0$. The conditions of Theorem 2.1 are satisfied and thus $\nu \geqq \mu$.

REFERENCES

[1] KARLIN, S. (1968). *Total Positivity* **1**. Stanford Univ. Press.

[2] KARLIN, S. and SCHUMAKER, L. (1967). The fundamental theorem of algebra for Tchebycheffian monosplines, *J. Analyse Math.* **20** 233–270.

[3] KARLIN, S. and STUDDEN, W. J. (1966a). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.

[4] KARLIN, S. and STUDDEN, W. J. (1966b). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley, New York.

[5] KARLIN, S. and ZIEGLER, Z. (1966). Tchebycheffian spline functions. *SIAM J. Appl. Math.* **3** 515–543.

[6] KIEFER, J. (1959). Optimum experimental designs. *J. Roy. Statist. Soc.* **21** 272–319.

[7] VANARMAN, D. (1968). Classification of experimental designs relative to polynomial spline regression functions. Mimeograph Series No. 166, Department of Statistics, Purdue Univ.