This research was supported by National Institutes of Health, Division of General Medical Sciences Grant No. 5 R01 GM 12868-05.

THE ADMISSIBILITY OF A RESPONSE SURFACE DESIGN

By

Lawrence L. Kupper

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 660

January 1970

ABSTRACT

The usual concept of the admissibility of an experimental design (e.g., see [6], pp. 800-812) is based solely on variance considerations and no allowance is made for the possible occurrence of bias errors resulting from the use of an incorrect model. This limitation is not entirely desirable in a response surface setting where a graduating function $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ such as a polynomial will always fail to represent exactly a response surface $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$. In fact, several pairs of authors such as Box and Draper ([1], [2]) and Draper and Lawrence ([3], [4]) have demonstrated in typical situations the overriding importance of bias considerations when selecting a design to minimize expected mean-square error $E(\hat{y}(\underset{\sim}{x};\underset{\sim}{b})-\eta(\underset{\sim}{x};\underset{\sim}{\beta}))^2$ averaged over a compact set X.

Motivated by their findings, this author first provides a generalization of a result on the minimization of integrated squared bias which was given in [1] and then uses it along with some ideas from decision theory to develop the notions of variance (V-), bias (B-) and mean-square (MS-) admissibility, which have particular appeal in a response surface framework. A theorem is proved characterizing a meaningful situation in which a design can be said to possess one of these types of admissibility.

1. Introduction. Let $f'(\underset{\sim}{x}) = (f_1(\underset{\sim}{x}), f_2(\underset{\sim}{x}), \ldots, f_k(\underset{\sim}{x}))$ be a vector of k real-valued functions defined on a given space X. In most applications, X is taken to be compact and the elements of $f(\underset{\sim}{x})$ taken to be continuous (so that suprema of certain functions are attained, insuring that optimum designs exist) and also linearly independent (so that trivial redundancies are avoided). For each "level" $\underset{\sim}{x} \in X$, an experiment can be performed whose outcome is a random variable $y(\underset{\sim}{x})$, where Var $y(\underset{\sim}{x}) = \sigma^2$. It is further assumed that $y(\underset{\sim}{x})$ has an expected value of the explicit form

$$(1.1) \qquad E\, y(\underset{\sim}{x}) = \textstyle\sum_{j=1}^{k} \beta_j f_j(\underset{\sim}{x}) = f'(\underset{\sim}{x})\underset{\sim}{\beta},$$

and that the "regression functions" $f_1, f_2, \ldots, f_k$ are known to the experimenter while the elements of the parameter vector $\underset{\sim}{\beta}' = (\beta_1, \beta_2, \ldots, \beta_k)$ are unknowns to be estimated on the basis of a finite number N of uncorrelated observations $\{y(\underset{\sim}{x}_i)\}_{i=1}^{N}$.

An exact experimental design corresponds to a probability measure δ on X concentrating positive weights $w_1, w_2, \ldots, w_p$ at the distinct points $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_p$, respectively, such that $w_i N = n_i$, $i=1,2,\ldots,p$, are integers, and $\sum_{i=1}^{p} w_i = 1$. Thus, the measure (or design) δ specifies the different points at which experiments take place, namely the $\{\underset{\sim}{x}_i\}_{i=1}^{p}$, and the number of experiments at each point, namely n_i at $\underset{\sim}{x}_i$. The $(k \times k)$ matrix

$$\frac{1}{N} X'X = \frac{1}{N} \textstyle\sum_{i=1}^{p} n_i f(\underset{\sim}{x}_i) f'(\underset{\sim}{x}_i),$$

where $X' = (f(\underset{\sim}{x}_1), f(\underset{\sim}{x}_2), \ldots, f(\underset{\sim}{x}_N))$, is called the information matrix of the exact design δ. The theory involved in the construction of exact experimental designs will be referred to as the <u>exact theory</u>.

More generally, if $\Delta = \{\delta\colon \delta$ is an arbitrary probability measure on the Borel sets $\mathcal{B}$ of $\mathcal{X}$, where $\mathcal{B}$ includes all one-point sets$\}$, then for each $\delta\varepsilon\Delta$ let us write

$$m_{ij}(\delta) = \int_{\mathcal{X}} f_i(\underset{\sim}{x}) f_j(\underset{\sim}{x}) d\delta(\underset{\sim}{x}), \tag{1.2}$$

and then define

$$M(\delta) = ((m_{ij}(\delta)))_{i,j=1}^{k}. \tag{1.3}$$

Thus, if δ is an exact design, $X'X = NM(\delta)$.

The following lemma, taken from Karlin and Studden [6], records several important properties of the information matrices $M(\delta)$.

<u>Lemma 1.1</u>. Let $M(\delta)$ be defined as in (1.2) and (1.3). Then,

(i) for each $\delta\varepsilon\Delta$, $M(\delta)$ is positive semi-definite;

(ii) $|M(\delta)| = 0$ if $p<k$;

(iii) the family of matrices $M(\delta)$, $\delta\varepsilon\Delta$, is a convex compact set;

(iv) for each $\delta\varepsilon\Delta$, the matrix $M(\delta)$ can be written as

$$\textstyle\sum_{i=1}^{p} w_i f(\underset{\sim}{x}_i) f'(\underset{\sim}{x}_i), \text{ where } p \le 1+\tfrac{1}{2}k(k+1).$$

The most important consequence of this lemma is contained in part (iv) which permits us to restrict attention to measures concentrated on a finite set of points when working with the matrices $M(\delta)$. Thus, for all practical purposes, Δ can be taken to be the set of all <u>discrete</u> probability measures on $\mathcal{X}$.

Many optimal design criteria are based on choosing a $\delta^*\varepsilon\Delta$ to minimize some real functional of $M(\delta)$, say $|M^{-1}(\delta)|$ or tr $M^{-1}(\delta)$, and, even though it can happen that δ^* takes on values other than multiples of $1/N$ and thus does not correspond to an exact design, it is worthwhile to consider this general approach to the construction of optimal designs, calling it the approximate theory and calling any measure $\delta\varepsilon\Delta$ an approximate design. The justification and convenience in allowing this greater generality in the choice of measures δ is that, in certain cases, it permits one to give a complete characterization of various optimal designs which is relevant for all N.

2. A Result on the Minimization of Integrated Squared Bias.

Let $\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1) = \underset{\sim}{f}'_{(1)}(\underset{\sim}{x})\underset{\sim}{b}_1$ and $Ey(\underset{\sim}{x}) = \eta(\underset{\sim}{x};\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2) = \underset{\sim}{f}'_{(1)}(\underset{\sim}{x})\underset{\sim}{\beta}_1 + \underset{\sim}{f}'_{(2)}(\underset{\sim}{x})\underset{\sim}{\beta}_2$ denote the forms of the estimated and true responses at $\underset{\sim}{x}\varepsilon X$, respectively; the vectors $f'_{(1)}(\underset{\sim}{x}) = (f_1(\underset{\sim}{x}),f_2(\underset{\sim}{x}),\ldots,f_{k_1}(\underset{\sim}{x}))$ and $f'_{(2)}(\underset{\sim}{x}) = (f_{k_1+1}(\underset{\sim}{x}),f_{k_1+2}(\underset{\sim}{x}),\ldots,f_{k_2}(\underset{\sim}{x}))$ are assumed to be known while the parameter vectors $\underset{\sim}{\beta}'_1 = (\beta_1,\underset{\sim}{\beta}_2,\ldots,\underset{\sim}{\beta}_{k_1})$ and $\underset{\sim}{\beta}'_2 = (\beta_{k_1+1},\beta_{k_1+2},\ldots,\beta_{k_2})$ are unknown, where $k_2>k_1$.

With observations at $N = \sum_{i=1}^{p} n_i$ points in X, we define the matrices $X'_1 = (f_{(1)}(\underset{\sim}{x}_1),f_{(1)}(\underset{\sim}{x}_2),\ldots,f_{(1)}(\underset{\sim}{x}_N))$ and $X'_2 = (f_{(2)}(\underset{\sim}{x}_1),f_{(2)}(\underset{\sim}{x}_2),\ldots,f_{(2)}(\underset{\sim}{x}_N))$, and we further assume that X_1 is of full rank. Then, if we agree to take $\underset{\sim}{b}_1$ as the standard least squares estimate of the form $(X'_1X_1)^{-1}X'_1\underset{\sim}{y}$, where $\underset{\sim}{y}' = (y(\underset{\sim}{x}_1),y(\underset{\sim}{x}_2),\ldots,y(\underset{\sim}{x}_N))$ is the vector of observations, it follows that $E\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1) = f'_{(1)}(\underset{\sim}{x})\underset{\sim}{\beta}_1 + f'_{(1)}(\underset{\sim}{x})A\underset{\sim}{\beta}_2$, the matrix $A = (X'_1X_1)^{-1}X'_1X_2$ being commonly referred to as the "alias matrix", where $X'_1X_1 = \sum_{i=1}^{p} n_i f_{(1)}(\underset{\sim}{x}_i)f'_{(1)}(\underset{\sim}{x}_i)$ and $X'_1X_2 = \sum_{i=1}^{p} n_i f_{(1)}(\underset{\sim}{x}_i)f'_{(2)}(\underset{\sim}{x}_i)$. So, from these results, it is clear that we can write the "squared bias" $(E\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1)-\eta(\underset{\sim}{x};\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2))^2$ in the form $[(f'_{(1)}(\underset{\sim}{x})A-f'_{(2)}(\underset{\sim}{x}))\underset{\sim}{\beta}_2]^2$, which is a function of the choice of the exact design δ only through the matrix A.

Now, since A can be written in the suggestive form $N(X_1'X_1)^{-1}\frac{1}{N}X_1'X_2$, it is meaningful to work in the more general setting of the approximate theory and to consider minimizing the integrated squared bias $B_\delta(w;\underset{\sim}{\beta}_2) = \int_{\mathcal{X}} B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)dw(\underset{\sim}{x})$ by choice of $\delta\varepsilon\Delta$, where the "weighting function" $w\varepsilon\mathcal{W}_{\mathcal{X}}$ = {w: w is a probability measure on $\mathcal{X}$} and where

$$B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2) = [(f'_{(1)}(\underset{\sim}{x})A(\delta)-f'_{(2)}(\underset{\sim}{x}))\underset{\sim}{\beta}_2]^2, \tag{2.1}$$

$$A(\delta) = M_{11}^{-1}(\delta)M_{12}(\delta), \tag{2.2}$$

$$M_{11}(\delta) = ((m_{ij}(\delta)))_{i,j=1}^{k_1} \tag{2.3}$$

$$M_{12}(\delta) = ((m_{ij}(\delta)))_{i=1,\ j=k_1+1}^{k_1 \quad k_2}, \tag{2.4}$$

the general expression for $m_{ij}(\delta)$ being given by (1.2). Note that we are taking $M_{11}(\delta)$ to be non-singular.

So, proceeding in this way, we have

$$B_\delta(w;\underset{\sim}{\beta}_2) = \int_{\mathcal{X}}\underset{\sim}{\beta}_2'(A'(\delta)f_{(1)}(\underset{\sim}{x})-f_{(2)}(\underset{\sim}{x}))(f'_{(1)}(\underset{\sim}{x})A(\delta)-f'_{(2)}(\underset{\sim}{x}))\underset{\sim}{\beta}_2\ dw(\underset{\sim}{x}) = \underset{\sim}{\beta}_2'\Gamma\underset{\sim}{\beta}_2,$$

where $\Gamma = A'(\delta)\Lambda_{11}A(\delta)-\Lambda_{12}'A(\delta)-A'(\delta)\Lambda_{12}+\Lambda_{22}$, with $\Lambda_{11} = \int_{\mathcal{X}} f_{(1)}(\underset{\sim}{x})f'_{(1)}(\underset{\sim}{x})dw(\underset{\sim}{x})$, $\Lambda_{12} = \int_{\mathcal{X}} f_{(1)}(\underset{\sim}{x})f'_{(2)}(\underset{\sim}{x})\ dw(\underset{\sim}{x})$, and $\Lambda_{22} = \int_{\mathcal{X}} f_{(2)}(\underset{\sim}{x})f'_{(2)}(\underset{\sim}{x})\ dw(\underset{\sim}{x})$. Now, assuming that Λ_{11} is non-singular, we can write $\Gamma = (\Lambda_{22}-\Lambda_{12}'\Lambda_{11}^{-1}\Lambda_{12}) + (A(\delta)-\Lambda_{11}^{-1}\Lambda_{12})'\Lambda_{11}(A(\delta)-\Lambda_{11}^{-1}\Lambda_{12}) = \Gamma_1+\Gamma_2$. Since

$$0 \le \int_{\mathcal{X}}[(f'_{(1)}(\underset{\sim}{x}),f'_{(2)}(\underset{\sim}{x}))\underset{\sim}{\alpha}]^2\ dw(\underset{\sim}{x}) = \underset{\sim}{\alpha}'\begin{pmatrix}\Lambda_{11} & \Lambda_{12}\\ \Lambda_{12}' & \Lambda_{22}\end{pmatrix}\underset{\sim}{\alpha},$$

it follows that Λ_{11} and $\begin{pmatrix}\Lambda_{11} & \Lambda_{12}\\ \Lambda_{12}' & \Lambda_{22}\end{pmatrix}$ are both positive semi-definite. Finally,

$$\begin{pmatrix} \Lambda_{11} & 0 \\ 0 & \Gamma_1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ -\Lambda_{12}'\Lambda_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{pmatrix} \begin{pmatrix} I & -\Lambda_{11}^{-1}\Lambda_{12} \\ 0 & I \end{pmatrix}$$

$$= T' \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{pmatrix} T,$$

which shows that Γ_1 is also positive semi-definite.

Thus, we have the following general result: no matter what the value of $\underset{\sim}{\beta}_2$, $B_\delta(w;\ \underset{\sim}{\beta}_2)$ is minimized by choice of $\delta\varepsilon\Delta$ when $A(\delta) = \Lambda_{11}^{-1}\Lambda_{12}$, and, in particular, when $M_{11}(\delta) = \Lambda_{11}$ and $M_{12}(\delta) = \Lambda_{12}$. In fact, from the structure of the pairs of matrices $M_{11}(\delta)$, Λ_{11} and $M_{12}(\delta)$, Λ_{12}, it follows directly that these latter two matrix equalities are satisfied by choice of $\delta\varepsilon\Delta$ when δ is taken to be the weighting function w itself. Since, in general, w is not discrete (e.g., it is chosen to be uniform over X in many instances), it is necessary to find a measure (design) concentrated only on a finite set of points in X which is equivalent to w in the sense that the pair of matrices Λ_{11} and Λ_{12} are the same for both w and its discrete counterpart.

This result follows as a direct generalization of work by Box and Draper in [1], who restrict their attention to the situation where the fitted and true models $\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1)$ and $\eta(\underset{\sim}{x};\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2)$ are polynomials of degree d_1 and d_2 ($>d_1$), respectively, and where the associated weighting function w is always uniform on X. The new treatment given here allows one to consider more general types of regression functions for both the fitted and true linear models and to characterize the associated optimum design with respect to any $w\varepsilon\mathcal{W}_X$ as an optimal <u>approximate</u> design. This author has done some work concerning the use of partial sums of Fourier series and spherical harmonics as response surface graduating functions and, in doing so, has found an important use for this generalized version of the Box-Draper result.

3. V-, B-, and MS-admissibility. The approximate theory approach to the optimal design of experiments as described in Section 1 characterizes an optimal design $\delta^* \varepsilon \Delta$ as one which minimizes some appropriate real functional of the information matrix $M(\delta)$ defined by the relations (1.2) and (1.3). In this light, the standard concepts that have so far been developed concerning the admissibility of an experimental design (see [6], pp. 808-812) have been based on the following definition.

Def. 3.1. A design $\delta \varepsilon \Delta$ is said to be admissible if there does not exist a design $\delta' \varepsilon \Delta$ such that $M(\delta') \geq M(\delta)$, where this inequality signifies that the matrix $M(\delta')-M(\delta)$ is positive semi-definite and $M(\delta') \neq M(\delta)$.

It is important to realize that this notion of admissibility is based solely on variance considerations and implicit in Def. 3.1 is the assumption that the model (1.1) exactly describes the true response at every point $\underset{\sim}{x} \varepsilon X$; in other words, no allowance is made for the possible occurrence of bias errors resulting from the use of an incorrect model. However, such an assumption is not entirely realistic in a response surface setting where a graduating function $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ such as a polynomial will always fail, at least to some extent, to represent a response surface $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$. In fact, several pairs of authors such as Box and Draper ([1], [2]) and Draper and Lawrence ([3], [4]) have demonstrated in typical situations the overriding importance of bias considerations when selecting a response surface design to minimize integrated mean-square error $\int_X E(\hat{y}(\underset{\sim}{x};\underset{\sim}{b}) - (\underset{\sim}{x};\underset{\sim}{\beta}))^2 dw(\underset{\sim}{x})$. (Their results will be discussed in more detail later.)

Motivated by these considerations, we now proceed to define the new concepts of V-, B-, and MS-admissibility, which have particular appeal in a response surface framework. In the notation of Section 2, let $e(\underset{\sim}{x},\delta)$ denote any of the following three functions:

$$V_\delta(\underset{\sim}{x}) = f'_{(1)}(\underset{\sim}{x})M_{11}^{-1}(\delta)f_{(1)}(\underset{\sim}{x}); \tag{3.1}$$

$$B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2); \tag{3.2}$$

and

$$MS_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2) = V_\delta(\underset{\sim}{x})+B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2), \tag{3.3}$$

where $B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)$, $A(\delta)$, $M_{11}(\delta)$ and $M_{12}(\delta)$ are given by (2.1)-(2.4).

Then, we have

Def. 3.2. A design $\delta\varepsilon\Delta$ is said to be variance (V-), bias (B-), or mean-square (MS-) admissible, depending on whether $e(\underset{\sim}{x},\delta)$ is equal to $V_\delta(\underset{\sim}{x})$, $B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)$, or $MS_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)$, respectively, if there does not exist a design $\delta'\varepsilon\Delta$ such that $e(\underset{\sim}{x},\delta') \leq e(\underset{\sim}{x},\delta)$ for every $\underset{\sim}{x}\varepsilon X$ and $e(\underset{\sim}{x}_0,\delta') < e(\underset{\sim}{x}_0,\delta)$ for some $\underset{\sim}{x}_0\varepsilon X$.

By noting that the expected mean-square error $E(\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1)-\eta(\underset{\sim}{x};\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2))^2$ is expressible as the sum of two distinct error terms, namely $E(\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1)-E\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1))^2 =$ $\mathrm{Var}\ \hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1)$ and $(E\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1)-\eta(\underset{\sim}{x};\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2))^2$, which, for an exact design δ, can be written (again, see Section 2) as $\frac{\sigma^2}{N} V_\delta(\underset{\sim}{x}) = f'_{(1)}(\underset{\sim}{x})(X_1'X_1)^{-1}f_{(1)}(\underset{\sim}{x})\sigma^2$ and $B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2) = [(f'_{(1)}(\underset{\sim}{x})(X_1'X_1)^{-1}X_1'X_2-f'_{(2)}(\underset{\sim}{x}))\underset{\sim}{\beta}_2]^2$, respectively, we can see that the properties of V-, B-, and MS- admissibility are certainly desirable ones for a response surface design to possess. In fact, one should feel somewhat disappointed in learning that a response surface design which minimized some meaningful function like maximum or integrated expected mean-square error, say, was not MS- admissible. Unfortunately, such a situation is not out of the realm of possibility; for example, a design which minimizes maximum squared bias is not necessarily B- admissible.

The following theorem, a take-off on a result in [5], p. 62, attempts to shed some light on this problem by characterizing an important situation in which a response surface design can be said to possess one of the types of admissibility introduced in Def. 3.2.

First, a preliminary definition.

<u>Def. 3.3</u>. Let X be a metric space with metric d. A point $\underset{\sim}{x}_o \varepsilon X$ is said to be in the <u>support</u> of a weighting function $w \varepsilon \mathcal{W}_X$ if, for every $\varepsilon > 0$, the neighborhood $N_\varepsilon(\underset{\sim}{x}_o) = \{\underset{\sim}{x} \varepsilon X: d(\underset{\sim}{x}, \underset{\sim}{x}_o) < \varepsilon\}$ has "positive weight" in the sense that $\int_{N_\varepsilon(\underset{\sim}{x}_o)} dw(\underset{\sim}{x}) > 0$.

Then, we have

<u>Theorem 3.1</u>. Let X be a metric space with metric d and assume that the functions (3.1), (3.2), and (3.3) are each continuous in $\underset{\sim}{x}$ for all $\delta \varepsilon \Delta$. Then, a design $\delta^* \varepsilon \Delta$ is V-, B-, or MS-admissible, depending on whether $e(\underset{\sim}{x}, \delta)$ has the form (3.1), (3.2), or (3.3), respectively, if δ^* minimizes $e(w^*, \delta) = \int_X e(\underset{\sim}{x}, \delta) dw^*(\underset{\sim}{x})$ with respect to some $w^* \varepsilon \mathcal{W}_X$ and if the support of w^* is X itself.

<u>Proof</u>. Assume that δ^* is not admissible in the sense of Def. 3.2. Then, there must exist a $\delta' \varepsilon \Delta$ for which $e(\underset{\sim}{x}, \delta') \leq e(\underset{\sim}{x}, \delta^*)$ for all $\underset{\sim}{x} \varepsilon X$ and $\eta = e(\underset{\sim}{x}_o, \delta^*) - e(\underset{\sim}{x}_o, \delta') > 0$ for some $\underset{\sim}{x}_o \varepsilon X$. Because $e(\underset{\sim}{x}, \delta)$ is continuous in $\underset{\sim}{x}$ for each $\delta \varepsilon \Delta$, there exists an $\varepsilon > 0$ such that $|e(\underset{\sim}{x}, \delta^*) - e(\underset{\sim}{x}_o, \delta^*)| \leq \eta/4$ and $|e(\underset{\sim}{x}, \delta') - e(\underset{\sim}{x}_o, \delta')| \leq \eta/4$ if $\underset{\sim}{x} \varepsilon N_\varepsilon(\underset{\sim}{x}_o)$. Thus, for $\underset{\sim}{x} \varepsilon N_\varepsilon(\underset{\sim}{x}_o)$, $e(\underset{\sim}{x}, \delta^*) - e(\underset{\sim}{x}, \delta') \geq (e(\underset{\sim}{x}_o, \delta^*) - \eta/4) - (e(\underset{\sim}{x}_o, \delta') + \eta/4) = e(\underset{\sim}{x}_o, \delta^*) - e(\underset{\sim}{x}_o, \delta') - \eta/2 = \eta/2 > 0$. Hence, we have $e(w^*, \delta^*) - e(w^*, \delta') = \int_X (e(\underset{\sim}{x}, \delta^*) - e(\underset{\sim}{x}, \delta')) dw^*(\underset{\sim}{x})$ $\geq \int_{N_\varepsilon(\underset{\sim}{x}_o)} (e(\underset{\sim}{x}, \delta^*) - e(\underset{\sim}{x}, \delta')) dw^*(\underset{\sim}{x}) \geq \frac{\eta}{2} \int_{N_\varepsilon(\underset{\sim}{x}_o)} dw^*(\underset{\sim}{x})$. But, since $\underset{\sim}{x}_o$ is in the support

of w^*, the last expression above is strictly greater than zero. This contradicts the assumption that δ^* minimizes $e(w^*,\delta)$ and completes the proof.

It is clear from the structure of the functions (3.1)-(3.3) that the continuity assumptions in Theorem 3.1 are not the least bit restrictive, and, in fact, obviously hold for all the standard regression models (e.g., polynomials and various trigonometric functions).

Theorem 3.1 provides incentive for constructing and using response surface designs which minimize a given measure of error between the fitted and true models that is averaged rather than, say, maximized over X; in particular, it lends support to the work of the pairs of authors mentioned earlier, who, in the framework of the exact theory, have constructed response surface designs for polynomial regression which minimize expected mean-square error or squared bias averaged with respect to uniform weighting functions over the k-dimensional hypersphere in [1] and [2] and over the three-dimensional simplex (equilateral triangle) in [3], and averaged with respect to a symmetric multivariate distribution weight function over k-dimensional Euclidean space in [4].

Finally, it is important to mention the fact the admissibility in the sense of Def. 3.1 is "stronger" than V- admissibility since a design $\delta\epsilon\Delta$ which is admissible in the former sense is clearly V- admissible, while the converse is not necessarily true.

REFERENCES

[1] BOX, G. E. P. and DRAPER, N. R. (1959). A basis for the selection of a response surface design. J. Amer. Statist. Assoc. 54, 622-654.

[2] BOX, G. E. P. and DRAPER, N. R. (1963). The choice of a second order rotatable design. Biometrika 50, 335-352.

[3] DRAPER, N. R. and LAWRENCE, W. E. (1965). Mixture designs for three factors. J. Roy. Statist. Soc. Ser. B. 27, 450-465.

[4] DRAPER, N. R. and LAWRENCE, W. E. (1967). Sequential designs for spherical weight functions. Technometrics 9, 517-529.

[5] FERGUSON, THOMAS S. (1967). Mathematical Statistics: A Decision Theoretic Approach. Academic Press, New York.

[6] KARLIN, SAMUEL and STUDDEN, WILLIAM J. (1966). Optimal experimental designs. Ann. Math. Statist. 37, 783-815.