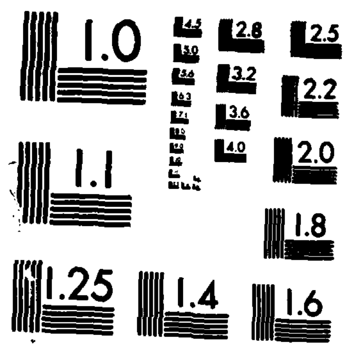
AD-A172 007 INFLUENTIAL NONNEGLIGIBLE PARAMETERS UNDER THE SEARCH LINEAR MODEL(U) CALIFORNIA UNIV RIVERSIDE DEPT OF STATISTICS S GHOSH 25 APR 86 TR-141 AFOSR-TR-86-0733 1/1

UNCLASSIFIED AFOSR-86-0048 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(2)

AFOSR-TR- 86-0738

INFLUENTIAL NONNEGLIGIBLE PARAMETERS UNDER THE SEARCH LINEAR MODEL*

BY

Subir Ghosh

University of California, Riverside

Summary

In this paper some results useful in detecting the Influential Nonnegligible parameters under the search linear model are presented. An estimator of the number of nonnegligible parameters which are significant and influential is also given.

AMS 1970 Subject Classifications: Primary and Secondary: 62J05, 62K05, 62K15, 62K99

Keywords and Phrases: Factorial Experiments, Influential Parameters, Linear Models, Residuals, Search, Treatment.

*The work of the author is sponsored by the Air Force Office of Scientific Research under Grant AFOSR-86-0048.

1. Introduction

Consider a general factorial experiment with the design consisting of t treatments and corresponding to the uth treatment there are n_u (≥ 1) observations and $\sum_{u=1}^{t} n_u = N$. Let y_{uv} be the observation corresponding to the uth replication of the vth treatment and $\bar{y}_u$ be the mean of all observations corresponding to the uth treatment. The model for this experiment is

$$\begin{aligned} E(\underline{y}) &= X_1\,\underline{\beta}_1 + X_2\,\underline{\beta}_2, \\ V(\underline{y}) &= \sigma^2 I, \\ \text{Rank } X_1 &= \nu_1, \end{aligned} \tag{1}$$

where $\underline{\beta}_1(\nu_1 x1)$ is a vector of specified lower order interactions and $\underline{\beta}_2(\nu_2 x1)$ is a vector of some or all of the higher order interactions, $X_1(N x \nu_1)$ and $X_2(N x \nu_2)$ are known matrices. It is known that K (very small compared to ν_2) elements of $\underline{\beta}_2$ are nonzero and the other are zero; however the value of K and the nonzero elements of $\underline{\beta}_2$ are unknown. The problem is to search the nonzero elements of $\underline{\beta}_2$ and draw inferences on them in addition to the elements of $\underline{\beta}_1$. Such a model is called the search linear model and was introduced in Srivastava (1975). Suppose K_1 is an initial guess on K. Note the three possibilities $K_1 > K$, $K_1 = K$ and $K_1 < k$. We consider $\binom{\nu_2}{K_1}$ models

$$\begin{aligned} E(\underline{y}) &= X_1\,\underline{\beta}_1 + X_2^{(i)}\underline{\beta}_2^{(i)}, \quad i=1,\ldots\binom{\nu_2}{K_1}, \\ V(\underline{y}) &= \sigma^2 I \\ \text{Rank}\left[X_1, X_2^{(i)}\right] &= \nu_1 + K_1, \end{aligned} \tag{2}$$

where $X_2^{(i)}(N\mathrm{x}K_1)$ is a submatrix of X_2 and $\underline{\beta}_2^{(i)}(K_1\mathrm{x}1)$ is a subvector of $\underline{\beta}_2$. It can be seen from Srivastava (1975) that we in fact need Rank $[X_1, X_2^{(i)}, X_2^{(i')}] = (\nu_1 + 2K_1)$, for all $i \neq i'$. This implies that $N \geq (\nu_1 + 2K_1)$. In case $K_1 = K$, one of $\binom{\nu_2}{K_1}$ models is the correct model. If $K_1 > K$, then $\binom{\nu_2-K}{K_1-K}$ models out of $\binom{\nu_2}{K_1}$ models include the true model as a submodel in the expectation forms of the models. The methods discussed in this paper will not only identify K nonzero parameters but also find how many of them have significant effects and, finally, rank the significant nonegligible parameters in the order of their influence on the fitted values. In case $K_1 < K$, the methods will identify from K_1 parameters the parameters which are significant and influential. We also propose an estimator of K in the Section 3.

In some industrial experiments, it is often easy to find replications ($n_u \geq 1$) in observations corresponding to a particular (the uth) treatments, see Taguchi and Wu (1985). There are also situations in industrial experiments where it is impossible to get replication in observations for a treatment, see Daniel (1976) and Box and Meyer (1985). The methods discussed in this paper consider both situations. In all Taguchi design methods, the higher order interactions (2-factor and higher order in most plans) are assumed to be zero. A few of those higher order interactions may be nonnegligible, significant and influential. The use of the search linear models may be a potential tool in improving upon the Taguchi design methods.

2. Influential Nonnegligible Parameters.

Let $Z_1^{(i)}$ $((N-\nu_1-K_1) \times N)$ be such that Rank $Z_1^{(i)} = (N-\nu_1-K_1)$, $Z_1^{(i)}Z_1^{(i)'} = I$ and $Z_1^{(i)}[X_1, X_2^{(i)}] = 0$. Let $Z^{(i)}(K_1 \times N)$ be such that Rank $\begin{bmatrix} Z_1^{(i)} \\ Z^{(i)} \end{bmatrix} = (N-\nu_1)$, $Z^{(i)}Z^{(i)'} = I$, $Z^{(i)}Z_1^{(i)'} = 0$ and $Z^{(i)}X_1 = 0$. It can be seen that under the ith model in (2), the minimum variance unbiased estimator (MVUE) of $\underline{\beta}_2^{(i)}$ is

$$\hat{\underline{\beta}}_2^{(i)} = \left(Z^{(i)}X_2^{(i)}\right)^{-1} Z^{(i)}\underline{y}. \tag{3}$$

In fact we can write $Z^{(i)'} = P_1X_2^{(i)}D^{(i)}$, where $D^{(i)}$ is a nonsingular (and triangular) matrix so that $Z^{(i)}Z^{(i)'} = I$ and $P_1 = I - X_1(X_1'X_1)^{-1}X_1'$. From the ith model in (2), the MVUE for $\underline{\beta}_1$ is

$$\hat{\underline{\beta}}_1^{(i)} = (X_1'X_1)^{-1}X_1'\underline{y} - (X_1'X_1)^{-1}X_1'X_2^{(i)}\hat{\underline{\beta}}_2^{(i)}. \tag{4}$$

The fitted value of $\underline{y}$ from the ith model in (2) is

$$\hat{\underline{y}}^{(i)} = X_1\hat{\underline{\beta}}_1^{(i)} + X_2^{(i)}\hat{\underline{\beta}}_2^{(i)}. \tag{5}$$

The residuals from the ith model in (2) are

$$\begin{aligned} \underline{R}^{(i)} &= \underline{y} - \hat{\underline{y}}^{(i)} = P_1\left(\underline{y}-X_2^{(i)}\hat{\underline{\beta}}_2^{(i)}\right) \\ &= P_1\left[I-X_2^{(i)}\left(X_2^{(i)'}P_1X_2^{(i)}\right)^{-1}X_2^{(i)'}\right]P_1\underline{y}\ . \end{aligned} \tag{6}$$

The sum of squares due to error under the ith $\left(i = 1,\ldots,\binom{\nu_2}{K_1}\right)$ model in (2) is

$$SSE^{(i)} = \underline{R}^{(i)'}\underline{R}^{(i)} = \underline{y}'Z_1^{(i)'}Z_1^{(i)}\underline{y}. \tag{7}$$

The residuals under the model (1), when $\underline{\beta}_2 = \underline{o}$, are

$$\underline{R}^{(o)} = \underline{y}-\hat{\underline{y}}^{(o)} = P_1\underline{y}. \tag{8}$$

It can be seen that

$$P_1 = Z_1^{(i)'}Z_1^{(i)} + Z^{(i)'}Z^{(i)}. \tag{9}$$

Therefore, for $i = 1,\ldots,\binom{\nu_2}{K_1}$,

$$SSE^{(o)} = \underline{R}^{(o)'}\underline{R}^{(o)} = SSE^{(i)} + \underline{y}'Z^{(i)'}Z^{(i)}\underline{y}. \tag{10}$$

For $i = 1,\ldots,\binom{\nu_2}{K_1}$, we define

$$F^{(i)} = \frac{\underline{y}'Z^{(i)'}Z^{(i)}\underline{y}/K_1}{SSE^{(i)}/(N-\nu_1-K_1)}. \tag{11}$$

Let $\hat{y}_u^{(i)}$ be the fitted value of the observation corresponding to the uth ($u = 1,\ldots,w$) treatment under the ith model in (2). We write the sum of squares due to lack of fit as

$$SSLOF^{(i)} = \sum_{u=1}^{w} n_u\,(\bar{y}_u-\hat{y}_u^{(i)})^2, \tag{12}$$

and the sum of squares due to pure error as

$$SSPE = \sum_{u=1}^{w}\ \sum_{v=1}^{n_u} (y_{uv}-\bar{y}_u)^2. \tag{13}$$

For $i = 1,\ldots,\binom{\nu_2}{K_1}$, we define

$$F_{LOF}^{(i)} = \frac{SSLOF^{(i)}/(w-\nu_1-K_1)}{SSPE/(N-w)}. \tag{14}$$

<u>Theorem 1</u>. For $\ell \in \{1,\ldots,\binom{\nu_2}{K_1}\}$, the following statements are equivalent.

(a) $SSE^{(\ell)}$ is a minimum,

(b) $F^{(\ell)}$ is a maximum,

(c) $SSLOF^{(\ell)}$ is a minimum,

(d) $F_{LOF}^{(\ell)}$ is a minimum,

(e) The Euclidean distance between $\underline{\hat{y}}^{(\ell)}$ and $\underline{\hat{y}}^{(o)}$ is a maximum,

(f) The square of the (sample) simple correlation coefficient between the elements of $\underline{R}^{(\ell)}$ and $\underline{R}^{(o)}$ is a minimum.

<u>Proof</u>. We have from (10) and (11) that

$$\left(\frac{K_1}{(N-\nu_1-K_1)}\, F^{(i)}\right)+1 = \frac{SSE^{(o)}}{SSE^{(i)}}\ .$$

Noting that the numerator on the RHS of the above expression does not depend on i, we get the equivalence of (a) and (b). Again,

$$SSE^{(i)} = SSPE + SSLOF^{(i)},$$

and SSPE does not depend on i. Therefore (a) and (c) are equivalent. From (14), the equivalence of (c) and (d) is clear. From (3), (6), (8) and (9), it follow that

$$\begin{aligned}
\underline{y}'Z^{(i)'}Z^{(i)}\underline{y} &= \underline{\hat{\beta}_2}^{(i)'}X_2^{(i)'}Z^{(i)'}Z^{(i)}X_2^{(i)}\ \underline{\hat{\beta}_2}^{(i)} \\
&= \underline{\hat{\beta}_2}^{(i)'}X_2^{(i)'}\left[P_1-Z_1^{(i)'}Z_1^{(i)}\right]\ X_2^{(i)}\underline{\hat{\beta}_2}^{(i)} \\
&= \underline{\hat{\beta}_2}^{(i)'}X_2^{(i)'}P_1X_2^{(i)}\underline{\hat{\beta}_2}^{(i)} \qquad (15) \\
&= (\underline{R}^{(i)}-\underline{R}^{(o)})'\ (\underline{R}^{(i)}-\underline{R}^{(o)}) \\
&= (-\underline{\hat{y}}^{(i)}+\underline{\hat{y}}^{(o)})'\ (-\underline{\hat{y}}^{(i)}+\underline{\hat{y}}^{(o)}).
\end{aligned}$$

The equivalence of (a) and (e) is now easy to see from (10) and (15). It follows from (10) and (15) that $\underline{R}^{(i)'}\underline{R}^{(i)} = \underline{R}^{(i)'}\underline{R}^{(o)}$. We thus have

$$\frac{SSE^{(i)}}{\underline{R}^{(o)'}\underline{R}^{(o)}} = \frac{\underline{R}^{(i)'}\underline{R}^{(i)}}{\underline{R}^{(o)'}\underline{R}^{(o)}} = \frac{\left(\underline{R}^{(i)'}\underline{R}^{(o)}\right)^2}{\left(\underline{R}^{(i)'}\underline{R}^{(i)}\right)\left(\underline{R}^{(o)'}\underline{R}^{(o)}\right)} \tag{16}$$

= the square of the (sample) Simple correlation Coefficient between $\underline{R}^{(i)}$ and $\underline{R}^{(o)}$.

The equivalence of (a) and (f) is now clear from (16). This completes the proof of the theorem.

Propostion 1. Under the ith model in (2),

$$Z^{(i)}\underline{R}^{(i)} = \underline{o}. \tag{17}$$

Proof. It follows from (3) and (5) that

$$Z^{(i)}\underline{Y} = Z^{(i)}X_2^{(i)}\hat{\underline{\beta}}_2^{(i)} = Z^{(i)}\hat{\underline{Y}}^{(i)}.$$

This completes the proof.

We have

$$V\left(\underline{R}^{(i)}\right) = \sigma^2 P_i \left[I - X_2^{(i)}\left(X_2^{(i)'}P_i X_2^{(i)}\right)^{-1} X_2^{(i)'}\right] P_i. \tag{18}$$

The residual in $\underline{R}^{(i)}$ are correlated and the question may be asked about the appropriateness in combining the elements of $\underline{R}^{(i)}$ in $SSE^{(i)}$. If we take the transformed residuals as $Z_i^{(i)}\underline{R}^{(i)}$, we then have

$$E\left(Z_i^{(i)}\underline{R}^{(i)}\right) = \underline{o} \text{ and } V\left(Z_i^{(i)}\underline{R}^{(i)}\right) = \sigma^2 I. \tag{19}$$

The sum of squares of these transformed residuals is $\underline{R}^{(i)'}Z_i^{(i)'}Z_i^{(i)}\underline{R}^{(i)}$.

<u>Proposition 2</u>. For $i = 1,\ldots,\binom{\nu_2}{K_1}$,

$$SSE^{(i)} = \underline{R}^{(i)'} Z_1^{(i)'} Z_1^{(i)} \underline{R}^{(i)}. \tag{20}$$

<u>Proof</u>. We write the RHS using (9) as

$$\underline{R}^{(i)'} Z_1^{(i)'} Z_1^{(i)} \underline{R}^{(i)} = \underline{R}^{(i)'} P_1 \underline{R}^{(i)} - \underline{R}^{(i)'} Z^{(i)'} Z^{(i)} \underline{R}^{(i)}. \tag{21}$$

It can be checked that $P_1\underline{R}^{(i)} = \underline{R}^{(i)}$. By using the Proposition 1, the rest of the proof is clear. This completes the proof.

Proposition 2 thus supports the use of $SSE^{(i)}$. Theorem 1 gives various interpretations of a search procedure, discussed in Srivastava (1975), of selecting $\underline{\beta}_2^{(\ell)}$ as the influential set of K_1 nonnegligible parameters.

We now denote

$$\begin{aligned}
\underline{\beta}_2^{(i)'} &= \left[\beta_{21}^{(i)},\ldots,\beta_{2j}^{(i)},\ldots,\beta_{2K_1}^{(i)}\right], \\
X_2^{(i)} &= \left[\underline{X}_{21}^{(i)},\ldots,\underline{X}_{2j}^{(i)},\ldots,\underline{X}_{2K_1}^{(i)}\right], \\
X_2^{(ij)} &= \text{the matrix obtained from } X_2^{(i)} \text{ by} \\
&\quad\ \text{deleting the jth column of } \underline{X}_{2j}^{(i)}. \\
X_{12}^{(ij)} &= \left[X_1,\ X_2{}^{(ij)}\right], \\
P_{12}^{(ij)} &= I - X_{12}^{(ij)}\left(X_{12}^{(ij)'}X_{12}^{(ij)}\right)^{-1} X_{12}^{(ij)'}, \\
\underline{z}_{1j}^{(i)} &= \frac{P_{12}^{(ij)}\underline{X}_{2j}^{(i)}}{\sqrt{\underline{X}_{2j}^{(i)'}P_{12}^{(ij)}\underline{X}_{2j}^{(i)}}}, \\
Z_o^{(i)'} &= \left[\underline{z}_{11}^{(i)},\ldots,\underline{z}_{1j}^{(i)},\ldots,\underline{z}_{1K_1}^{(i)}\right].
\end{aligned} \tag{22}$$

It can be seen that

$$\text{Rank} \begin{bmatrix} Z_i^{(i)} \\ \underline{Z}_{1j}^{(i)'} \end{bmatrix} = (N - \nu_i - K_i + 1),$$

$$\underline{Z}_{1j}^{(i)'} \underline{Z}_{1j}^{(i)} = 1, \ \underline{Z}_{1j}^{(i)'} Z_i^{(i)} = \underline{0}', \tag{23}$$

$$\begin{bmatrix} Z_i^{(i)} \\ \underline{Z}_{1j}^{(i)'} \end{bmatrix} X_{12}^{(ij)} = \underline{0}, \ \text{Rank } Z_0^{(i)} = K_i.$$

There exists a nonsingular (triangular) matrix $D_0^{(i)}$ such that

$$Z^{(i)} = D_0^{(i)} Z_0^{(i)}. \tag{24}$$

From (3) and (24), we have

$$\underline{\hat{\beta}}_2^{(i)} = (Z_0^{(i)} X_2^{(i)})^{-1} Z_0^{(i)} \underline{y}. \tag{25}$$

Now

$$Z_0^{(i)} X_2^{(i)} = \text{diag } (\underline{Z}_{11}^{(i)'} X_{21}^{(i)} \ ,\ldots, \underline{Z}_{1j}^{(i)'} X_{2j}^{(i)} ,\ldots, \underline{Z}_{1K}^{(i)'} \underline{X}_{2K}^{(i)}) \tag{26}$$

is a diagonal matrix. Thus

$$\hat{\beta}_{2j}^{(i)} = \frac{\underline{Z}_{1j}^{(i)'} \underline{y}}{\underline{Z}_{1j}^{(i)'} \underline{X}_{2j}^{(i)}} . \tag{27}$$

Let $\underline{R}^{(ij)}$, $i = 1,\ldots, \binom{\nu_2}{K_i}$, $j = 1,\ldots,K_i$, be the residuals obtained from ith model in (2) assuming $\beta_{2j}^{(i)} = 0$. Then the sum of squares due to error is

$$SSE^{(ij)} = \underline{R}^{(ij)'} \underline{R}^{(ij)} = (\underline{Z}_{1j}^{(i)'} \underline{y})^2 + SSE^{(i)}. \tag{28}$$

We now define, for $i = 1,\ldots,\binom{\nu_2}{K_i}$ and $j = 1,\ldots,K_i$,

$$t^{(ij)} = \frac{\underline{z}_{1j}^{(i)'}\ \underline{y}}{\sqrt{\dfrac{SSE^{(i)}}{(N-\nu_1-K_1)}}} \ . \tag{29}$$

Proposition 3. For a fixed ℓ in $\{1,\ldots,\binom{\nu_2}{K_1}\}$ and an m in $\{1,\ldots,K_1\}$, the following statements are equivalent.

(a) $SSE^{(\ell m)}$ is a minimum,

(b) $t^{(\ell m)}$ is a maximum.

Proof. The proof can be easily seen from (28) and (29).

In the set $\underline{\beta}_2^{(\ell)}$ of influential nonnegligible parameters, $\beta_{2m}^{(\ell)}$ is the most influential nonnegligible parameters. The influential nonnegligible parameters may or may not have significant effects on observations.

3. Influential Significant Nonnegligible Parameters

We now assume the normality in (2) and therefore for

$i = 1,\ldots,\binom{\nu_2}{K_1}$, $\underline{y} \underbrace{\quad}_{}^{\text{independent}} N(X_1\underline{\beta}_1 + X_2^{(i)}\underline{\beta}_2^{(i)}, \sigma^2 I)$.

Under the null hypothesis H_0: $\underline{\beta}_2^{(i)} = \underline{0}$, $F^{(i)}$ has the central F distribution with $(K_1,\ N - \nu_1 - K_1)$ d.f. and under the null hypothesis H_0: $\beta_{2j}^{(i)} = 0$, $t^{(ij)}$ has the central t distribution with $(N - \nu_1 - K_1)$ d.f. We now present a further development of a procedure suggested in Srivastava (1975).

Case I. If $\max_i F^{(i)} \leq F_{\alpha;K_1,N-\nu_1-K_1}$, we then conclude that there is no significant nonnegligible parameter. ($F_{\alpha;K_1,N-\nu_1-K_1}$ is the upper α percent point of the central F distribution with $(K_1,N-\nu_1-K_1)$d.f.).

Case II. Suppose for $i = i_1,\ldots,i_s$, we have $F^{(i)} > F_{\alpha;K_1,N-\nu_1-K_1}$. We denote for $j = 1,\ldots,\nu_2$,

$$\partial_j = \text{the number of } i \text{ in } \{i_1,\ldots,i_s\} \text{ for which } |t_{ij}| > t_{\frac{\alpha}{2},\ N-\nu_1-K_1}.$$

Note that $0 \leq \partial_j \leq s$. We now arrange ∂_j's in decreasing order of magnitude and write $\partial_{(1)} \geq \partial_{(2)} \geq \cdots \geq \partial_{(\nu_2)}$. If there are at least K_1 nonzero $\partial_{(j)}$'s, we select the influential significant parameters as $\beta_{(1)},\ldots,\beta_{(K_1)}$, otherwise we pick the influential $\beta_{(j)}$'s corresponding to nonzero $\partial_{(j)}$'s (Note that the number of influential parameters is then less than K_1). The parameter $\beta_{(1)}$ is the most influential significant nonnegligible parameter. An estimator of the unknown K is $\hat{K}$ = the number of nonzero ∂_j'S, $j = 1,\ldots,\nu_2$.

4. Miscellaneous Results

4.a. Let us denote the unknown nonzero elements of $\underline{\beta}_2$ in (1) by $\underline{\beta}_{2c}$ (Kx1) and the zero elements of $\underline{\beta}_2$ by $\underline{\beta}_{2d}((\nu_2-K)X1)$, the corresponding columns in X_2 matrix are X_{2c} and X_{2d}. The unknown true expectation form of (1) is thus

$$E(\underline{y}) = X_1\underline{\beta}_1 + X_{2c}\underline{\beta}_{2c}. \tag{30}$$

The expectation form of the ith model in (2) can be written as

$$E(\underline{y}) = X_1\underline{\beta}_1 + X_{2c}^{(i)}\underline{\beta}_{2c}^{(i)} + X_{2d}^{(i)}\underline{\beta}_{2d}^{(i)}, \tag{31}$$

where $X_{2c}^{(i)}(N x\gamma_i)$ is a submatrix of X_{2c}, $X_{2d}^{(i)}(Nx(K_1-\gamma_i))$ is a submatrix of X_{2d}, $\underline{\beta}_{2c}^{(i)}(\gamma_i x1)$ is a subvector of $\underline{\beta}_{2c}$ and $\underline{\beta}_{2d}^{(i)}$ $((K_1-\gamma_i)X1)$ is a subvector of $\underline{\beta}_{2d}$. Let $\underline{\beta}_{2c}^{*(i)}$ is the vector of elements in $\underline{\beta}_{2c}$ which are not in $\underline{\beta}_{2c}^{(i)}$ and $X_{2c}^{*(i)}$ is the matrix whose columns are in X_{2c} but not in $X_{2c}^{(i)}$. The following result, a counterpart of the result in (10) for the population, can be verified very easily.

<u>Proposition 4</u>. Under (30),

$$\begin{aligned} E\left(SSE^{(i)}\right) &= E\left(SSLOF^{(i)}\right) + \sigma^2(N-w) \\ &= \sigma^2(N-\nu_1-K_1) + \underline{\beta}_{2c}' X_{2c}' Z_1^{(i)'} Z_1^{(i)} X_{2c}\underline{\beta}_{2c} \\ &= \sigma^2(N-\nu_1-K_1) + \underline{\beta}_{2c}^{*(i)'} X_{2c}^{*(i)'} Z_1^{(i)'} Z_1^{(i)} X_{2c}^{*(i)} \underline{\beta}_{2c}^{*(i)} \\ &= E\left(SSE^{(0)}\right) - \left[\sigma^2 K_1 + \underline{\beta}_{2c}' X_{2c}' Z^{(i)'} Z^{(i)} X_{2c}\underline{\beta}_{2c}\right] . \end{aligned} \tag{32}$$

4.b. The model obtained from (2)

$$E(Z^{(i)}\underline{y}) = Z^{(i)}X_2^{(i)}\underline{\beta}_2^{(i)},$$

$$V(Z^{(i)}\underline{y}) = \sigma^2 I,$$

is called the pure search model (Srivastava (1976)). In fact, Srivastava (1976) considered a special form of $Z^{(i)}$.

4.c. The influential nonnegligible parameter may depend on noise, i.e., a parameter may be influential under one noise but may not be influential under another noise.

4.d. The replicated observations will surely improve the chances of detecting the correct influential nonnegligible parameters.

4.e. In presence of outliers in observations, one may combine residuals with unequal weights, or in other words, may use transformed residuals (see, Cook and Weisberg (1982)).

4.e.1. An example of transformed residuals is the vector $M^{(i)}\underline{R}^{(i)}$ where $M^{(i)}(N\text{x}N)$ is a diagonal matrix whose uth diagonal element is $\left(1/\sqrt{m_{uu}^{(i)}}\right)$ with $m_{uu}^{(i)}$ being the uth diagonal element of $\sigma^{-2}V(\underline{R}^{(i)})$.

4.e.2. Suppose the underlying design is robust against the unavailability of any single observation [see, Ghosh (1980)] in the sense that the estimation of $\underline{\beta}_1$ and $\underline{\beta}_2^{(i)}$ is possible under (2) when any single observation is unavailable. We find the predicted value of the uth observation from the remaining (N-1) observations (i.e., by deleting the uth observation). The difference between the uth observation and its predicted value is called the uth predicted residual (using the idea of cross validation). It can be verified algebraically that the vector of predicted residuals is $[M^{(i)}]^2 \underline{R}^{(i)}$. The predicted residual sum of squares (PRESS) from the ith model under (2) is

$$\text{PRESS}^{(i)} = \underline{R}^{(i)\prime} [M^{(i)}]^4 \underline{R}^{(i)}$$

In presence of outliers, one may take $\text{PRESS}^{(i)}$ as an alternative to $\text{SSE}^{(i)}$.

BIBLIOGRAPHY

Box, G. E. P. and Meyer, R. D. (1985) Studies in quality improvement I: dispersion effects from fractional designs. Tech. Report #2796. Mathematics Research Center, University of Wisconsin - Madison.

Cook, R. D. and Weisberg, S (1982). Residuals and Influence in Regression. Chapman and Hall, New York,

Daniel, C. (1976). Applications of Statistics to Industrial Experimentation. J. Wiley & Sons, New York.

Ghosh, S. (1980). On robustness of optional balanced resolution V plans. J. of Statistical Planning and Inference, 4, 313-319.

Ghosh, S. (1983). Influence observations in view of design and inference. Commun. Statist. - Theor. Meth., 12(14), 1675-1983.

Srivastava, J. N. (1975). Designs for searching nonnegligible effects. In A Survey of Statistical Design and Linear Models (J. N. Srivastava, Ed.), pp. 507-519. North-Holland, Amsterdam.

Srivastava, J. N. (1976). Some further theory of search linear models. In Cotributions to Applied Statistics, Experimentia: Suppl. (W.J. Ziegler, Ed.), Vol. 22. Birkhäuser, Basel.

Srivastava, J. N. and Mallenby, D. W. (1985). On a decision rule using dichotomies for identifying the nonnegligible parameter in certain linear models. J. of Multivariate Analysis, 16, 318-334.

Taguchi, G. and Wu, Y. (1985). Introduction to Off-Line Quality Control. Central Japan Quality Control Association, Tokyo.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS: A172 007

2a. SECURITY CLASSIFICATION AUTHORITY: NA

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: NA

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release, Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S): Technical Report No. 141

5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 86-0733

6a. NAME OF PERFORMING ORGANIZATION: University of California, Riverside

6b. OFFICE SYMBOL (If applicable):

7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM

6c. ADDRESS (City, State and ZIP Code): DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, CA 92521

7b. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR

8b. OFFICE SYMBOL (If applicable): NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-86-0048

8c. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, DC

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
6.1102F	2304	A5	

11. TITLE (Include Security Classification): Influential Non-negligible Parameters Under The Search Linear Model

12. PERSONAL AUTHOR(S): SUBIR GHOSH

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
INTERIM	FROM ______ TO ______	86-04-25	15

16. SUPPLEMENTARY NOTATION: SUBMITTED IN COMMUNICATIONS IN STATISTICS

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Factorial Experiments, Influential Parameters, Linear Models, Residuals, Search, Treatment

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

In this paper some results useful in detecting the Influential Nonnegligible parameters under the search linear model are presented. An estimator of the number of nonnegligible parameters which are significant and influential is also given.

86 9 19 074

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: MAJOR BRIAN W. WOODRUFF

22b. TELEPHONE NUMBER (Include Area Code): (202)767-5027

22c. OFFICE SYMBOL: NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

END
DTIC
11-86