CONSTRUCTION OF SOME ODD RESOLUTION SATURATED FRACTIONAL REPLICATES AND SOME POSSIBLE VALUES OF THE DETERMINANTS OF SEMINORMALIZED (0,1) TREATMENT COMBINATION MATRICES

by

Walter T. Federer

BU-436-M

November, 1972

ABSTRACT

A simple and straightforward method of constructing saturated fractional replicate plans of odd resolution for any factorial is presented. It is shown that the saturated mean and main effect plan constructed in this manner from the 2^n factorial, has the largest variance-covariance matrix possible for all such plans, thus resulting in a plan of minimal variance optimality. It is further shown how to construct saturated mean and main effect plans which take on specified absolute values of the determinant of the seminormalized (0,1) treatment combination matrix for fractional replicates from the 2^n factorial.

CONSTRUCTION OF SOME ODD RESOLUTION SATURATED FRACTIONAL REPLICATES AND SOME POSSIBLE VALUES OF THE DETERMINANTS OF SEMINORMALIZED (0,1) TREATMENT COMBINATION MATRICES

by

Walter T. Federer

BU-436-M

November, 1972

1. INTRODUCTION

A simple and straightforward method of constructing saturated fractional replicate plans of odd resolution for any factorial is presented. It is shown that the saturated mean and main effect plan constructed in this manner from the 2^n factorial, has the largest variance-covariance matrix possible for all such plans, thus resulting in a plan of minimal variance optimality. It is further shown how to construct saturated mean and main effect plans which take on specified absolute values of the determinant of the seminormalized (0,1) treatment combination matrix for fractional replicates from the 2^n factorial.

BU-436-M in the Mimeograph Series of the Biometrics Unit, Cornell University. Research partially supported under an NIH Research Grant No. 5-R01-GM-05900.

2. A METHOD OF CONSTRUCTING SATURATED FRACTIONAL REPLICATES OF ODD RESOLUTION FROM ANY FACTORIAL

A very simple method of obtaining saturated fractional replicate plans from any factorial of n factors, say A_i, i = 1,2,...,n, with any number of levels, say k_i, is illustrated through an example from a 2 x 3 x 4 factorial of factors A_1, A_2, and A_3 with $k_1 = 2$, $k_2 = 3$, and $k_3 = 4$. Let the combination of the three factors be denoted by hij where, h = 0 or 1, i = 0,1, or 2, and j = 0, 1, 2, or 3. Note that the method does not depend upon equal spacings of the levels as the values for h, i, and j merely designate the level and not the actual amount. We shall write the factorial effects in the order of the number of factors involved, although this is arbitrary, and shall list the degrees of freedom (df) associated with each factorial effect. The example follows:

Effect	df	hij
mean	1	000
A_1	1	100
A_2	2	010 020
A_3	3	001 002 003
$A_1 \times A_2$	2	110 120
$A_1 \times A_3$	3	101 102 103
$A_2 \times A_3$	6	011 012 013 021 022 023
$A_1 \times A_2 \times A_3$	6	111 112 113 121 122 123

(000 through 003): mean, A_1, A_2, and A_3 = resolution III

(000 through 103): mean, A_1, A_2, A_3, $A_1 \times A_2$, and $A_1 \times A_3$

(000 through 023): mean, A_1, A_2, A_3, $A_1 \times A_2$, $A_1 \times A_3$ and $A_2 \times A_3$ = resolution V

(000 through 123): full factorial with all effects estimable

The number of observations added in each case to estimate another effect is equal to the number of degrees of freedom associated with that effect. Further, <u>any</u> set of effects and the corresponding observations, may be added in <u>any</u> order to obtain a saturated fractional replicate. For example, a saturated fractional replicate for the mean, A_1, A_2, A_3, and $A_1 \times A_2 \times A_3$ would result from taking the first seven and the last six combinations. Likewise, to estimate the mean and the two-factor interactions, a saturated fractional replicate results from taking the first and the eighth through the eighteenth combinations. It should also be noted that the entire effect, and not a subset, should be considered, and that this method for main effects is the familiar one-at-a-time method.

In general then, a saturated main effect plan (resolution III) may be constructed as follows:

Combination					Effect	df
0	0	0	$\cdots$	0	mean	1
1	0	0	$\cdots$	0		
2	0	0	$\cdots$	0	A_1	k_1-1
$\vdots$						
k_1-1	0	0	$\cdots$	0		
0	1	0	$\cdots$	0		
0	2	0	$\cdots$	0	A_2	k_2-1
$\vdots$						
0	k_2-1	0	$\cdots$	0		
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$		
0	0	0	$\cdots$	1		
0	0	0	$\cdots$	2	A_n	k_n-1
$\vdots$				$\vdots$		
0	0	0	$\cdots$	k_n-1		

This saturated mean and main effect fractional replicate requires $1 + \sum_{i=1}^{n} (k_i - 1)$ observations associated with the corresponding combinations. The proof that such observations result in a nonsingular information matrix is straightforward. For each factor, say A_i, set up the actual levels of the factor in relation to the single degree of freedom contrasts, for example as in polynomial regression, as follows:

$$\begin{bmatrix} 1 & X_{i0} & X_{i0}^2 & \cdots & X_{i,0}^{k_i-1} \\ 1 & X_{i1} & X_{i1}^2 & & X_{i,1}^{k_i-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{i,k_i-1} & X_{i,k_i-1}^2 & & X_{i,k_i-1}^{k_i-1} \end{bmatrix}$$

If one first subtracts the first row from the last k_i-1 rows, expands on the first column, and if one then shows that the remaining $(k_i-1) \times (k_i-1)$ cofactor is nonsingular, we note that the determinant of this matrix is always nonzero. This is true for each i and hence the entire matrix obtained for the above plan is nonsingular.

To augment the previous plan to estimate the $A_1 \times A_2$ interaction, simply add the following $(k_1-1)(k_2-1)$ combinations:

1	1	0	...	0
1	2	0	...	0
⋮				
1	k_2-1	0	...	0
2	1	0	...	0
2	2	0	...	0
⋮				⋮
2	k_2-1	0	...	0
⋮				⋮
k_1-1	1	0	...	0
⋮				⋮
k_1-1	k_2-1	0	...	0

To estimate all two-factor interactions, proceed in the above manner by taking all possible pairs of factors. The additional number of combinations is equal to $\sum_{i=1}^{n-1} \sum_{\substack{j=2 \\ i<j}}^{n} \left(k_i-1\right)\left(k_j-1\right)$. The resulting saturated fraction is of resolution V.

Continuing this procedure, any odd resolution saturated fractional replicate may be obtained.

3. OPTIMALITY PROPERTIES OF SATURATED MEAN AND MAIN EFFECT PLANS OBTAINED IN SECTION 2

In this section we show that the saturated main effect plans are the least variance optimal of all nonsingular plans that can be constructed from the 2^n factorial. We conjecture that these are also the worst possible plans from a variance-optimality criterion from any factorial.

Let the n + 1 set of combinations from a 2^n factorial for estimating the mean and the n main effects be $D_{(n+1) \times n}$. For the one-at-a-time plan of the previous section $D_{(n+1) \times n} = \begin{pmatrix} 0'_{1 \times n} \\ \hline I_{n \times n} \end{pmatrix}$ where $0'_{1 \times n}$ is a row vector of zeros and $I_{n \times n}$ is the n x n identity matrix. Then from Raktoe and Federer [1970], we note that $X^* = \left(\begin{array}{c|c} 1 & 0'_{1 \times n} \\ \hline 1_{n \times 1} & I_{n \times n} \end{array}\right)$ may be used instead of the semi-normalized X matrix of plus and minus ones to study optimality properties of X'X.

Theorem 3.1: <u>Of all nonsingular saturated main effect plans of n + 1 combinations from the 2^n factorial, the one-at-a-time plan $D_{(n + 1) \times n} = \begin{pmatrix} 0'_{1 \times n} \\ \hline I_{n \times n} \end{pmatrix}$ has the minimal absolute value of the determinant of $X^* = \left(\begin{array}{c|c} 1 & 0'_{1 \times n} \\ \hline 1_{n \times 1} & I_{n \times n} \end{array}\right)$, and also of $(X^*)'X^*$.</u>

<u>Proof</u>: From the definition of a determinant (e.g., see Ferrar [1941], page 8) it may be noted that the sum of all possible permutations of the products of the n + 1 coefficients, one from each column such that the permutation is of the n + 1 column numbers, must be zero or one. Hence, the sum can only be integer valued. Since the absolute value of determinant of X^* is one, this is the minimal value of the sum that is possible, since zero is excluded. Then, $||(X^*)'X^*||$ is

also equal to one and the variance-covariance matrix takes on its maximum value, resulting in the least optimal nonsingular plan possible for estimating the mean and the n main effects from a 2^n factorial.

It is not immediately clear to this writer how to proceed for the general factorial although it is intuitively clear that each level of a factor must be present and that this is the most unequal representation of levels of each factor possible in the sense of having the maximum number at the zero level and one at each of the other levels. It is conceivable that smaller absolute values of the determinant of X'X could be obtained from plans other than the one-at-a-time plan but this is not considered to be a very likely situation, and hence, it is conjectured that a theorem similar to Theorem 3.1 holds for all factorials.

One interesting item is that the main effect plan of n + 1 combinations is least-optimal, but yet as one proceeds to the full factorial, full information is obtained. This, then, raises the question about least-optimality for resolution V and higher odd resolution plans.

4. CONSTRUCTION OF PLANS FOR VARIOUS VALUES OF THE ABSOLUTE VALUE OF THE DETERMINANT OF THE MATRIX X^*

An important result for the 2^n factorial would be to obtain the possible values of the determinants of all $\binom{2^n}{n+1}$ possible plans for the 2^n factorial. From Paik and Federer [1970b] we note that $||X^*|| = 1$ for $n = 2$, $||X^*|| = 0,1,2$ for $n = 3$, $||X^*|| = 0,1,2,3$ for $n = 4$, and that these are all the values possible. Wells [1971], page 241, has shown that there are more than n values (absolute) of the determinant possible for $n = 5,6,7$. Hence, the question of the possible absolute values of determinant of X^* for general n is an unresolved question. If this were known, then the results of Paik and Federer [1970a, 1972], Raktoe and Federer [1971], and others could be utilized in determining the number of plans for each value of the determinant of X^*.

An alternative approach is to characterize classes of plans $D_{(n+1) \times n}$ such that $||X^*||$ takes on a class of values. We do this in the following theorem.

Theorem 4.1: Nonsingular saturated main effect plans of $n + 1$ combinations from the 2^n factorial may be constructed such that (i) $||X^*|| = 1,2,\ldots,n-1$; (ii) $||X^*|| = \prod_{i=1}^{c} a_{is}$ where a_{is} is a member of the set $\{1,2,\ldots,d_i-2\}$, $\sum_{i=1}^{c} d_i = n$, and $d_i \geq 4$.

Proof: The proof is by construction. Since $X^*_{(n+1)(n+1)} = \left(\begin{array}{c|c} 1 & 0'_{1 \times n} \\ \hline 1_{n \times 1} & I_{n \times n} \end{array}\right)$

has the minimal absolute value of the determinant, one should not have to change the structure appreciably to obtain larger values. If one simply replaces $0'_{1 \times n}$ by $1'_{1 \times n}$ which is a row vector of ones, then $||X^*|| = \left\|\begin{array}{c|c} 1 & 1'_{1 \times n} \\ \hline 1_{n \times 1} & I_{n \times n} \end{array}\right\| = n - 1.$

Likewise, if one substitutes $p(0 \leq p \leq n-2)$ zeros for p of the ones in $1'_{1 \times n}$, then the values of $||X^*|| = n - p - 1$. This completes the proof for (i).

For the proof of (ii), construct $X^* =$

$$\left[\begin{array}{c|cccc} 1 & & 0'_{1 \times n} & & \\ \hline & X^*_{d_1 \times d_1} & 0 & & 0 \\ \mathbf{1}_{n \times 1} & 0 & X^*_{d_2 \times d_2} & & 0 \\ & \vdots & & & \vdots \\ & 0 & 0 & & X^*_{d_c \times d_c} \end{array}\right]$$

Since we know from (i) that $\|X^*_{d_i \times d_i}\| = 1,2,\cdots,d_i-2$, the proof of (ii) follows immediately.

The proof of (ii) may be extended in that the set $\{1,2,\cdots,d_i-2\}$ may be enlarged when $\|X^*_{d_i \times d_i}\|$ can take on values other than those in the set given. Also, it would be interesting to note what values $\|X^*_{(n+1) \times (n+1)}\|$ can take on when $0'_{1 \times n}$ is replaced by $\mathbf{1}'_{1 \times n}$ and when $0 \leq p \leq n-2$ zeros replace ones in $\mathbf{1}'_{1 \times n}$. It is interesting to note the possible values of $\|X^*\|$ for a particular example. Let n = 20. Then

$$\|X^*_{21 \times 21}\| = \left\|\begin{array}{c|cc} 1 & 0'_{1 \times 20} & \\ \hline \mathbf{1}_{20 \times 1} & X^*_{10 \times 10} & 0 \\ & 0 & X^*_{10 \times 10} \end{array}\right\| = a_{1s}a_{2s} \text{ where } a_{is} = \{1,2,\cdots,8\}.$$

The largest value possible is 64. When $X^*_{9 \times 9}$ and $X^*_{11 \times 11}$ plans are used the maximum product is 63, and when $X^*_{8 \times 8}$ and $X^*_{12 \times 12}$ are used the maximum product is 60. However, when four $X^*_{5 \times 5}$ plans are inserted on the diagonals as in (ii), the maximum product is $3^4 = 81$. Again it should be emphasized that the set a_{is} may take on values other than those given. This would increase the number of values that $\|X^*_{21 \times 21}\|$ can take on in (ii).

In addition to extending the above for the 2^n factorial, extension to the s^n factorial as well as for the general $A_1 \times A_2 \times \ldots A_n$ factorial where the i^{th} factor is at k_i levels would be a worthwhile project both statistically and mathematically. Another extension is the construction of an X^* for the s^n factorial.

The following table gives absolute values of the determinant (M), number of plans by Wells [1971] (W) for $\binom{2^n}{n}$ plans and number of plans by Paik and Federer [1970b] (P) for $\binom{2^n}{n+1}$ plans, for use in extending (ii) of Theorem 4.1:

n = 3

M	W	P
0	27	6(2) = 12
1	28	7(8) = 56
2	1	1(2) = 2

n = 4

M	W	P
0	880	64(16) = 1024
1	835	(188)(16) = 3008
2	100	(20)(16) = 320
3	5	1(16) = 16

n = 5

M	W
0	97090
1	80856
2	20232
3	2412
4	726
5	60

n = 6

M	W
0	34923518
1	25666809
2	10746288
3	2135343
4	1163064
5	176701
6	129360
7	17885
8	13930
9	1470

n = 7

M	W
0	40885781314
1	26883246720
2	16511989560
3	4650079360
4	3511706880
5	744944448
6	833612648
7	161359296
8	208846176
9	57084608
10	42833560
11	9880640
12	17749760
13	2437120
14	2432640
15	806400
16	759360
17	80640
18	135240
19	0
20	26880
21	0
22	0
23	0
24	1920
25	0
26	0
27	0
28	0
29	0
30	0
31	0
32	30

5. RESULTS ON THE POSSIBLE RANKS OF AND BOUNDS ON TREATMENT COMBINATION MATRICES

If one writes the $(n+1) \times n$ treatment generator matrix as $\left(0_{n \times 1} \vdots D^{*\prime}_{n \times n}\right)'$, then $D^*_{n \times n}$ consists of all sets of n combinations from the 2^n factorial except the 0,0,$\cdots$,0 combination. Also, since determinants of $X^* = \begin{pmatrix} 1 & \vdots & 0'_{1 \times n} \\ 1_{n \times 1} & \vdots & D^*_{n \times n} \end{pmatrix}$ may be formulated in terms of $D^*_{n \times n}$, possible values for the rank of X^* may be approached through looking at possible ranks of $D^*_{n \times n}$. The following theorem is in this spirit.

Theorem 5.1: *The possible ranks of the matrix $D^*_{n \times n}$ for $2^{p-1} \le n \le 2^p - 1$ are p, p+1,$\cdots$,n.*

Proof: The proof for $n = 2^p - 1$ is by construction. Construct D^* equal to $\begin{pmatrix} I_{p \times p} & 0 \\ LI_{p \times p} & 0 \end{pmatrix}$ where $LI_{p \times p}$ is a $(2^p-1-p) \times p$ matrix whose rows are linear combinations of $I_{p \times p}$ and all distinct. Hence all rows in D^* are distinct combinations and the rank of D^* equals rank of $I_{p \times p}$ which is p. Now substitute one combination which is not a linear combination of the rows of $I_{p \times p}$ for one of the last 2^p-1-p rows of D^*; the rank will be p + 1. Continue this process until full rank, 2^p-1, is attained.

The proof for $2^{p-1} \le n < 2^p - 1$ is by contradiction. We first select $I_{(p-1) \times (p-1)}$. Note that there are too few rows to construct the remaining rows of D^* and that it is necessary to use $I_{p \times p}$ in order to construct the remaining rows of D^* as linear combinations of the first p rows.

Therefore, the possible ranks of $X^*_{(n+1) \times (n+1)}$ are p+1,p+2,$\cdots$,n+1.

An upper bound on the absolute value of the determinant of X* may be obtained from Hadamard's theorem as shown by Raktoe and Federer [1970] and is $2^{-n}(n+1)^{(n+1)/2}$. This upper bound may be improved slightly as is shown in the following theorem:

Theorem 5.2: An upper bound on the absolute value of the determinant of X* is:

$$||X^*|| \leq \text{integer part of } 2^{-n}(n+1)^{(n+1)/2} .$$

Proof: From the definition of a determinant only integer values are possible. Hence, the equality sign for the value $||X^*||$ can only occur when $2^{-n}(n + 1)^{(n+1)/2}$ is an integer. Since only integer values are permitted the largest value $||X^*||$ can take on is the integer part of $2^{-n}(n + 1)^{(n+1)/2}$.

The upper bounds are compared in the following table:

n	$2^{-n}(n+1)^{(n+1)/2}$	integer part of column 2	actual value obtained
2	1.299	1	1
3	2	2	2
4	3.494	3	3
5	6.75	6	5
6	14.181	14	9
7	32	32	32
8	76.887	76	?
9	195.3125	195	?
10	521.627	521	?
11	1,458	1,458	1,458
12	4,248.85	4,248	?
13	12,867.9	12,867	?
14	40,389.1	40,389	?
15	131,072	131,072	131,072

Note that for $4t \leq n < n+1 < n+2 \leq 4(t+1)$, $2t^{2t} \leq 2^{-n+1}n^{n/2}$ $< 2^{-n}(n+1)^{(n+1)/2} < (n+2)^{(n+2)/2} \leq 2(t+1)^{2t+2}$.

Theroem 5.3: *If $4t \le n+1 \le 4(t+1)$ and if a $4t \times 4t$ Hadamard matrix exists, a lower bound on the maximum absolute value of $X^*_{(n+1) \times (n+1)}$ is $2t^{2t}$ and an upper bound is the integer part of $2^{-n}(n+1)^{(n+1)/2}$. In addition, if $n \ge 4t$, the maximum value of $\|X^*_{n \times n}\|$ serves as a lower bound for the maximum value of $\|X^*_{(n+1) \times (n+1)}\|$.*

Proof: Let
$$X^*_{(n+1) \times (n+1)} = \left(\begin{array}{c|c} 1 & 0'_{n \times n} \\ \hline \mathbf{1}_{n\times 1} & D^*_{n \times n} \end{array}\right) = \left(\mathbf{1}_{(n+1) \times 1} \,\vdots\, D_{(n+1 \times n)}\right).$$

Let $D^*_{n \times n}$ be the plan providing the maximum absolute value of the determinant. Then expanding on the first row, we note that the $\|X^*_{(n+1) \times (n+1)}\|$ is equal to $\|D^*_{n \times n}\|$; and whether or not this value can be increased is determined by $D_{(n+1) \times n}$. Since we know that Hadamard matrices exist for all values of $4t$ up to 200 except 188 we can construct $D_{4t \times 4t}$ which achieves the maximum absolute value of $\|X^*_{(n+1) \times (n+1)}\|$ equal to $2t^{2t}$ for many values of $4t$.

6. LITERATURE CITED

Ferrar, W. L. [1941]. *Algebra. A text-book of determinants, matrices, and algebraic forms.* Oxford University Press, London, Edinburgh, Glasgow, New York, Toronto, Melbourne, Capetown, Bombay, Calcutta, Madras.

Paik, U. B. and Federer, W. T. [1970a]. A randomized procedure of saturated main effect fractional replicates. Annals of Mathematical Statistics 41: 369-375.

Paik, U. B. and Federer, W. T. [1970b]. On construction of fractional replicates and on aliasing schemes. Mathematical Research Center Technical Summary Report No. 1029, U. S. Army and the University of Wisconsin.

Paik, U. B. and Federer, W. T. [1972]. On a randomized procedure for saturated fractional replicates in a 2^n-factorial. Annals of Mathematical Statistics 43:1346-1351.

Raktoe, B. L. and Federer, W. T. [1970]. Characterization of optimal saturated main effect plans of the 2^n factorial. Annals of Mathematical Statistics 41:203-206.

Raktoe, B. L. and Federer, W. T. [1971]. On the number of generators of saturated main effect fractional replicates. Annals of Mathematical Statistics 42: 1758-1760.

Wells, M. B. [1971]. Elements of Combinatorial Computing. Pergamon Press, Oxford, New York, Toronto, Sydney, Braunschweig.