ON A RANDOMIZED PROCEDURE FOR SATURATED FRACTIONAL REPLICATES IN A 2^n-FACTORIAL

by

U. B. Paik and W. T. Federer

Korea University and Cornell University

Abstract

Paik and Federer [1970] presented a randomized procedure for non-orthogonal saturated main effect fractional replicates in a s^n-factorial and presented an unbiased estimator of the main effect parameter vector. However, the explicit expression of the variance of the estimator remained an unsolved problem. In this paper our attention is restricted to a 2^n-factorial, and the randomized procedure is extended to any preassigned parameters in a 2^n-factorial system. An explicit expression of the variances of unbiased estimators of the parameters is presented. Also, in a 2^n-factorial, some invariant properties of the information matrices and variances of the estimators in the randomized fractional replicates and a semi-invariant property of alias schemes of the fractional replicates are obtained.

Paper No. BU-333-M in the Biometrics Unit, October, 1970.

This research was supported by NIH GM-05900-12 grant at Cornell University.

ON A RANDOMIZED PROCEDURE FOR SATURATED FRACTIONAL REPLICATES IN A 2^n-FACTORIAL

by

U. B. Paik and W. T. Federer

Korea University and Cornell University

BU-333-M Oct. 1970

0. Introduction and summary. Paik and Federer [4] presented a randomized procedure for non-orthogonal saturated main effect fractional replicates in an s^n-factorial and presented an unbiased estimator of the main effect parameter vector. However, the explicit expression of the variance of the estimator remained an unsolved problem. In this paper our attention is restricted to a 2^n-factorial, and the randomized procedure is extended to any preassigned parameters in a 2^n-factorial system. An explicit expression of the variances of unbiased estimators of the parameters is presented. Also, in a 2^n-factorial, some invariant properties of the information matrices and variances of the estimators in the randomized fractional replicates and a semi-invariant property of alias schemes of the fractional replicates are obtained.

1. Basic notations and statistical model. In a 2^n-factorial system, the space of treatment combinations, Z, is represented by the set $Z = \{(i_1, i_2, \ldots, i_n): i_h = 0 \text{ or } 1 \text{ for all } h = 1,2,\ldots,n\}$ which contains 2^n points, say $N = 2^n$. A standard ordering of points in Z is given by the relationship between the coordinate of a point $z_v = (i_1, i_2, \ldots, i_n)$, $v = 0,1,\ldots,N-1$, and order subscript $v = \Sigma_{h=1}^n i_h 2^{n-h}$.

The addition operation + and multiplication operation $\cdot$ of any of two treatment combinations z_v and $z_{v'}$ are defined as addition and inner product of two row vectors of $z_v = (i_1, i_2, \ldots, i_n)$ and $z_{v'} = (i_1', i_2', \ldots, i_n')$, modulo 2, respectively. It follows immediately that the set Z is a group with respect to addition.

The expected value of the random vector $y(Z)$ associated with the space of treatment combinations Z is given by

$$E\left[y(Z)\right] = X\underline{B},$$

where X is a $2^n \times 2^n$ matrix with orthogonal column vectors such that $X'X = 2^n I$, $\underline{B}$ is the $N\times 1$ column vector of single degrees of freedom parameters $\beta_0, \beta_1, \ldots, \beta_{N-1}$ and $y(Z)$ is the $N\times 1$ column vector of observations, with covariance matrix $\sigma^2 I$. The parameters β_u have the usual interpretation of main effects and interactions of n factors. We further describe the structure of N parameters, β_u, $u = 0,1,\ldots,N-1$, by considering the space B of N points where $B = \{(\alpha_1,\alpha_2,\ldots,\alpha_n): \alpha_k = 0 \text{ or } 1$ for all $h = 1,2,\ldots,n\}$. The correspondence between the parameters and the points of B is given by the order relation specified by $u = \Sigma_{h=1}^{n}\alpha_h 2^{n-h}$. We also introduce addition and the inner product of any of two row vectors $\beta_u = (\alpha_1,\alpha_2,\ldots,\alpha_n)$ and $\beta_u' = (\alpha_1',\alpha_2',\ldots,\alpha_n')$ or between z_v and β_u. The unit element of this parameter group in addition, $\beta_0 = (0,0,\ldots,0)$, is the mean response of all the treatment combinations. The parameter point β_u in which the k^{th} position is 1 and all other positions are zero, corresponds to the k^{th} factor first degree main effect. Interactions correspond to points where coordinates are zero or non-zero with at least two non-zero coordinates. The matrix X can be defined as

$$X = X^{(2)} \otimes \ldots \otimes X^{(2)} ,$$

where $X^{(2)} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and $\otimes$ denotes Kronecker product.

Suppose that the vectors of $y(Z)$ and $\underline{B}$ are rearranged and partitioned as follows: $y(Z^*)' = (y(Z_p)', y(Z_{N-p})')$, $\underline{B}^{*'} = (\underline{B}_p', \underline{B}_{N-p}')$, where $y(Z_p)$ and $\underline{B}_p$ are $p\times 1$ observations and $p\times 1$ interesting parameter vectors, respectively, with the mean parameter as the first element of $\underline{B}_p$. We shall write $\underline{y}_p$ and $\underline{y}_{N-p}$ for $y(Z_p)$ and $y(Z_{N-p})$, respectively, and also we shall use a new notation for the rearranged parameters in vectors $\underline{B}_p$ and $\underline{B}_{N-p}$ such that $\underline{B}_p = (b_0, b_1, \ldots, b_{p-1})'$ and $\underline{B}_{N-p} = (b_1^*, b_2^*, \ldots, b_{N-p}^*)'$.

Consider the expression $E\left[y(Z)\right] = \left[X_1, X_2\right]\left[\underline{B}_p', \underline{B}_{N-p}'\right]$ where X_1 is an $N \times p$ matrix and X_2 is an $N \times (N-p)$ matrix. The matrix $\left[X_1, X_2\right]$ is obtained by rearranging the column order in X and the partitioning of that matrix. Since X is an $N \times N$ matrix with orthogonal column vectors, $r(X_1) = p$. Hence, there exists at least one non-singular $p \times p$ matrix X_{11} in the matrix X_1.

After rearranging the order of the elements in $y(Z)$ and the row order in $\left[X_1, X_2\right]$, respectively, we obtain the following expression:

$$E\left[\underline{y}_p\right] = \left[X_{11}, X_{12}\right]\left[\underline{B}_p', \underline{B}_{N-p}'\right]'$$

such that X_{11} is a non-singular $p \times p$ matrix, the observations in $\underline{y}_p$ yield a saturated fractional replicate for the given parameter vector $\underline{B}_p$.

Using the lease squares procedure we obtain the solution (Banerjee and Federer [1], Zacks [7]),

$$\hat{\underline{B}}_p^* = X_{11}^{-1}\underline{y}_p \ .$$

Hence, $X_{11}^{-1}\underline{y}_p$ is the best linear unbiased estimator of $\underline{B}_p + X_{11}^{-1}X_{12}\underline{B}_{N-p}$.

2. <u>A property of matrices X_{11} and X_{12}</u>. Let Z_p be a saturated fractional replicate plan given a parameter vector $\underline{B}_p$ represented by a submatrix of Z such as a $p \times n$ matrix in a 2^n-factorial and X_{11} by a $p \times p$ coefficient matrix of

$\underline{B}_p$ and X_{12} by a p×(N-p) coefficient matrix of $\underline{B}_{N-p}$ corresponding to the plan Z_p, and let $J(i_1,i_2,\ldots,i_n)$ be a p×n matrix such that

$$J(i_1,i_2,\ldots,i_n)=\begin{bmatrix} i_1,i_2,\ldots,i_n \\ \cdot\ \cdot\ \cdot\ \cdot \\ i_1,i_2,\ldots,i_n \end{bmatrix}$$

where $i_h = 0$ or 1 for all $h = 1,2,\ldots,n$ and $X_{11,v}$ and $X_{12,v}$ be a p×p coefficient matrix of $\underline{B}_p$ and a p×(N-p) coefficient matrix of $\underline{B}_{N-p}$ corresponding to the plan $Z_{p,v} = Z_p + J(i_1,i_2,\ldots,i_n) \bmod 2$, where the order subscript $v = \Sigma_{h=1}^{n} i_h 2^{n-h}$.

Let $G_{p,v}$ and $G_{N-p,v}$ be a p×p diagonal matrix and an (N-p) × (N-p) diagonal matrix, respectively, such that the $(1+u)^{th}$ diagonal element of $G_{p,v}$ is $d_v(u) = (-1)^{z_v b_u}$ and the r^{th} diagonal element of G_{N-p} is $d_v^*(r) = (-1)^{z_v b_r^*}$. Then, since each diagonal element of $G_{p,v}$ and $G_{N-p,v}$ is -1 or 1, $G_{p,v}^{-1} = G_{p,v}$ and $G_{N-p,v}^{-1} = G_{N-p,v}$ for all $v = 0, 1,\ldots,N-1$ and the lemma presented below may be easily verified. First however, note the following:

Suppose b_1 and b_2 are main effect parameters such that $b_1 = (1,0,\ldots,0)$ and $b_2 = (0,1,0,\ldots,0)$ and let $b_k = b_1 + b_2 = (1,1,0,\ldots,0)$, and suppose ξ_{1j}, ξ_{2j}, and ξ_{kj} are the j^{th} elements of the columns in $\left[X_{11},X_{12}\right]$ corresponding to parameters b_1, b_2 and b_k, respectively, then $\xi_{kj} = \xi_{1j}\xi_{2j}$. Consider $Z_{p,v} = Z_p + J(1,0,1,\ldots,1)$ and let $\xi_{1j,v}$, $\xi_{2j,v}$, and $\xi_{kj,v}$ be the j^{th} elements of the columns in $\left[X_{11,v}, X_{12,v}\right]$ corresponding to the parameters b_1, b_2, and b_k, respectively, then $\xi_{1j,v} = (-1)\xi_{1j} = \xi_{1j}(-1)^{(1,0,1,\ldots,1)(1,0,\ldots,0)}$, $\xi_{2j,v} = \xi_{2j} = \xi_{2j}(-1)^{(1,0,1,\ldots,1)(0,1,0,\ldots,0)}$, and $\xi_{kj,v} = \xi_{1j,v}\xi_{2j,v}$
$= \xi_{1j}\xi_{2j}(-1)^{(1,0,1,\ldots,1)(1,0,\ldots,0) + (1,0,1,\ldots,1)(0,1,0,\ldots,0)}$

$= \xi_{kj}(-1)^{(1,0,1,\ldots,1)(1,1,0,\ldots,0)}$.

<u>Lemma</u> (i) $X_{11,v} = X_{11}G_{p,v}$, $X_{12,v} = X_{12}G_{N-p,v}$.

(ii) $\sum_{v=0}^{N-1} d_v(u) = 0$ for all $u = 1,2,\ldots,p-1$, and $\sum_{v=0}^{N-1} d_v^*(r) = 0$ for all $r = 1,2,\ldots,N-p$.

(iii) $$\sum_{v=0}^{N-1} d_v(u)d_v(u') = \begin{cases} N & \text{if } u = u' \\ 0 & \text{if } u \neq u' \end{cases}$$

$$\sum_{v=0}^{N-1} d_v^*(r)d_v^*(r') = \begin{cases} N & \text{if } r = r' \\ 0 & \text{if } r \neq r' \end{cases}$$

(iv) $$\sum_{v=0}^{N-1} d_v(u)d_v^*(r) = 0$$

(v) $$\sum_{v=0}^{N-1} d_v(u)d_v(u')d_v^*(r)d^*(r') = \begin{cases} N & \text{if } u = u' \text{ and } r = r' \\ 0 & \text{if } u = u' \text{ but } r \neq r' \\ 0 & \text{if } u \neq u' \text{ but } r = r' \\ N & \text{if } u \neq u' \text{ and } r \neq r' \\ & \quad \text{but } b_u + b_{u'} + b_r^* + b_{r'}^* = b_0 \\ 0 & \text{if } u \neq u',\ r \neq r' \\ & \quad \text{and } b_u + b_{u'} + b_r^* + b_{r'}^* \neq b_0 . \end{cases}$$

Note that $d_v(u)d_v(u') = (-1)^{z_v(b_u + b_{u'})}$ and

$$d_v(u)d_v(u')d_v^*(r)d_v^*(r') = (-1)^{z_v(b_u + b_{u'} + b_r^* + b_{r'}^*)} .$$

3. An invariant property of $|X_{11}'X_{11}|$ and a semi-invariant property of $X_{11}'X_{11}$ and $X_{11}^{-1}X_{12}\underline{B}_{N-p}$. Paik and Federer [5] presented some patterns of $X_{11}^{-1}X_{12}$ in irregular fractional replicates in a 2^n-factorial as this gives the aliasing scheme for the fractional replicate. Now, we define a semi-invariant property of $X_{11}^{-1}X_{12}$ such that if the matrix $X_{11}^{-1}X_{12}$ remains unchanged except the sign of each element under the procedure $Z_p + J(i_1,i_2,\ldots,i_n)$ where $i_h = 0$ or 1 for all $h = 1,2,\ldots,n$, we say that the matrix $X_{11}^{-1}X_{12}$ is semi-invariant under such a procedure. Also, we define a notation abs(A) such that if $A = \|a_{ij}\|$, $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$, $\text{abs}(A) = \| \, |a_{ij}| \, \|$, $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$, where $|a_{ij}|$ denotes the absolute value of the element a_{ij}.

Since $X_{11,v} = X_{11}G_{p,v}$, $|G_{p,v}| = 1$, $X_{12,v} = X_{12}G_{N-p,v}$, and all elements of the diagonal matrices $G_{p,v}$ and $G_{N-p,v}$ are -1 or 1, the following theorem may be easily verified.

Theorem 1. If Z_p is a saturated fractional replicate plan in a 2^n-factorial system given $\underline{B}_p$, then $Z_{p,v}$ also is a saturated fractional replicate plan of $\underline{B}_p$ and $|X_{11,v}'X_{11,v}| = |X_{11}'X_{11}|$, $\text{abs}\left(X_{11,v}^{-1}X_{12,v}\right) = \text{abs}\left(X_{11}^{-1}X_{12}\right)$.

The meaning of this theorem is that if Z_p is not a subgroup of Z in a 2^n-factorial, $Z_p + J(i_1,i_2,\ldots,i_n)$, $i_h = 0$ or 1 for all $h = 1,2,\ldots,n$, may produce 2^n different saturated fractional replicate plans of $\underline{B}_p$, but the determinants of the information matrices have the same value. Furthermore, the information matrices and the aliasing matrices have the semi-invariant property.

4. A randomized procedure for saturated fractional replicates. Ehrenfeld and Zacks [2,3] presented randomized procedures for regular fractions, and Zacks [7] showed that an unbiased estimator of a given parameter vector in the saturated

fractional replicate case exists only if one randomizes over all possible designs of a certain structure. Paik and Federer [4] give a method similar to the Randomized Procedure I in the above papers and present an unbiased estimator of the main effect parameter vector for irregular saturated fractional replicates. As an extension of the authors' results, an unbiased estimator and the corresponding variances are given below for any saturated fractional replicate for the 2^n-factorial.

A saturated fractional replicate plan $Z_{p,v}$ of $\underline{B}_p$ in a 2^n-factorial is said to be <u>independent</u> of a saturated fractional replicate plan Z_p if $Z_{p,v}$ cannot be constructed by the procedure $Z_p + J(i_1, i_2, \ldots, i_n)$, $i_h = 0$ or 1 for all $h = 1,2,\ldots,n$. If $Z_{p,v}$ and Z_p are not independent then the plan $Z_{p,v}$ is an element of the set, $S(Z) = \{Z_p + J(i_1, i_2, \ldots, i_n) : i_h = 0 \text{ or } 1 \text{ for all } h = 1,2,\ldots,n\}$. The set $S(Z_p)$ is said to be the saturated fractional replicate plan set of $\underline{B}_p$ <u>generated</u> by Z_p. Paik and Federer [5] presented a complete list of the generators of the saturated main effect plans in the cases for 2^2, 2^3 and 2^4 factorials, and Raktoe and Federer [6] obtained a formula for the number of generators of saturated main effect fractional replicates for s^n-factorials.

Define a notation sq(A) such that if $A = \|a_{ij}\|$, $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$, then $\text{sq}(A) = \|a_{ij}^2\|$, $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$.

<u>Theorem 2</u>. <u>Suppose a saturated fractional replicate plan $Z_{p,v}$ of $\underline{B}_p$ is chosen at random from a set generated by a plan Z_p in a 2^n-factorial, then, given plan $Z_{p,v}$, the least squares estimator $\hat{B}^*_{p,v} = X^{-1}_{11,v}\underline{y}_p$ of $\underline{B}^*_{p,v} = \underline{B}_p + X^{-1}_{11,v}X_{12,v}\underline{B}_{N-p}$ is an unbiased estimator of $\underline{B}_p$ and</u>

$$\left(V(\hat{b}_0),\ V(\hat{b}_1),\ldots,V(\hat{b}_{p-1})\right)' = \underline{1}'\left[\text{sq}\left(X_{11}^{-1}\right)\right]\sigma^2 + \left[\text{sq}\left(X_{11}^{-1}X_{12}\right)\right]\left(b_1^{*2}, b_2^{*2}, \ldots, b_{N-p}^{*2}\right)',$$

<u>where $\underline{1}$ is a p×1 vector having all elements unity.</u>

Note that the variance vector $\left(V(\hat{b}_0),\ V(\hat{b}_1),\dots,V(\hat{b}_{p-1})\right)'$ is invariant under the procedure $Z_p + J(i_1,i_2,\dots,i_n)$, $i_h = 0$ or 1 for all $h = 1,2,\dots,n$.

Proof: Let $X_{11}^{-1}X_{12} = \|w_{st}\|$, $s = 1,2,\dots,p;\ t = 1,2,\dots,N-p$, then

$$E\hat{\underline{B}}^{*}_{p,v} = E_v\left\{E\left[\hat{\underline{B}}^{*}_{p,v} \mid X_{11,v}\right]\right\}$$

$$= E_v\left\{X_{11,v}^{-1}\left(X_{11,v}\underline{B}_p + X_{12,v}\underline{B}_{N-p}\right)\right\}$$

$$= \underline{B}_p + \left(E\left[X_{11,v}^{-1}X_{12,v}\right]\right)\underline{B}_{N-p}$$

$$= \underline{B}_p + \left(E_v\left[G_{p,v}\left(X_{11}^{-1}X_{12}\right)G_{N-p,v}\right)\underline{B}_{N-p}$$

$$= \underline{B}_p + \left[E_v\|d_v(s)d_v^{*}(t)w_{st}\|\right]\underline{B}_{N-p}, \text{ where } s = 1,2,\dots,p;\ t = 1,2,\dots,N-p,$$

$$= \underline{B}_p, \text{ by Lemma (iv).}$$

Next,

$$\mathrm{Cov}\left(\hat{\underline{B}}^{*}_{p,v}\right) = E_v\left\{\mathrm{Cov}\left(\hat{\underline{B}}^{*}_{p,v} \mid X_{11,v}\right)\right\} + \mathrm{Cov}_v\left\{E\left(\hat{\underline{B}}^{*}_{p,v} \mid X_{11,v}\right)\right\}$$

$$= E_v\left(X'_{11,v}X_{11,v}\right)^{-1}\sigma^2 + \mathrm{Cov}_v\left(\underline{B}^{*}_{p,v}\right).$$

Let $X_{11}^{-1} = \|x_{ij}\|$, $i = 0, 1,\dots,p-1;\ j = 0,1,\dots,p-1$, then

$$E_v\left(X'_{11,v}X_{11,v}\right)^{-1} = E_v\left(G_{p,v}X'_{11}X_{11}G_p\right)^{-1}$$

$$= E_v\left[G_{p,v}\left(X'_{11}X_{11}\right)^{-1}G_p\right]$$

$$= E_v\left[\left\|d_v(i)d_v(j)\sum_{k=0}^{p-1} x_{ki}x_{kj}\right\|\right], \text{ where } i,j = 0,1,\dots,p-$$

$$= \begin{bmatrix} \Sigma x_{k0}^2 & & 0 \\ & \ddots & \\ 0 & & \Sigma x_{k,p-1}^2 \end{bmatrix}.$$

$$\mathrm{Cov}_v\left(\underline{B}^*_{p,v}\right) = E_v\left(X^{-1}_{11,v}X_{12,v}\underline{B}_{N-p}\right)\left(X^{-1}_{11,v}X_{12,v}\underline{B}_{N-p}\right)'$$

$$= E_v\left(G_{p,v}X^{-1}_{11}X_{12}G_{N-p,v}\underline{B}_{N-p}\underline{B}'_{N-p}G_{N-p,v}X'_{12}X'^{-1}_{11}G_{p,v}\right)$$

$$= E_v\left[\left\| d_v(i)d_v(j)\sum_{t=1}^{N-p}\sum_{r=1}^{N-p} w_{it}w_{jr}d^*_v(r)d^*_v(t)b^*_r b^*_t\right\|\right],$$

where $i,j = 0,1,\ldots,p-1$,

$$= E_v\left[\left\| \sum_{t=1}^{N-p}\sum_{r=1}^{N-p} d_v(i)d_v(j)d^*_v(r)d^*_v(t)w_{it}w_{jr}b^*_r b^*_t\right\|\right],$$

where $i,j = 0,1,\ldots,p-1$.

Using the Lemma (v),

$$V_v\left(\hat{b}^*_1\right) = \sum_t w^2_{1t}b^{*2}_t,\ V_v\left(\hat{b}^*_2\right) = \sum_t w^2_{2t}b^{*2}_t,\ldots,\ V_v\left(\hat{b}^*_{N-p}\right) = \sum_t w^2_{N-p,t}b^{*2}_t .$$

Hence, we can write:

$$\left(V\left(\hat{b}_0\right),\ V\left(\hat{b}_1\right),\ldots,V\left(\hat{b}_{p-1}\right)\right)' = \mathbf{1}'\left[sq\left(X^{-1}_{11}\right)\right]\sigma^2 + \left[sq\left(X^{-1}_{11}X_{12}\right)\right]\left(b^{*2}_1, b^{*2}_2,\ldots,b^{*2}_{N-p}\right)'.$$

Note: (i) The variances of estimators are not the same value, i.e., they are dependent on the choice of a generator of a non-orthogonal fraction in a 2^n-factorial.

(ii) Since $d_v(i)d_v(j)d^*_v(r)d^*_v(t) = (-1)^{z_v(b_i + b_j + b^*_r + b^*_t)}$ and $b_i + b_j + b^*_r + b^*_t$ could be b_0, the estimator $\hat{b}_0, \hat{b}_1,\ldots,\hat{b}_{p-1}$ are not always uncorrelated, i.e., the off diagonal elements of $\mathrm{Cov}_v\left(\underline{B}^*_{p,v}\right)$ are not always zero and these are dependent on the choice of a generator of non-orthogonal fractions and choice of parameter vector $\underline{B}_p$ in a 2^n-factorial.

References

[1] Banerjee, K. S. and Federer, W. T. [1964], "Estimates of effects for fractional replicates." Ann. Math. Statist. 35, 711-715.

[2] Ehrenfeld, S. and Zacks, S. [1961], "Randomization and factorial experiments." Ann. Math. Statist. 32, 270-297.

[3] Ehrenfeld, S. and Zacks, S. [1967], "Testing hypothesis in randomized factorial experiments." Ann. Math. Statist. 38, 1494-1507.

[4] Paik, U. B. and Federer, W. T. [1970], "A randomized procedure of saturated main effect fractional replicates." Ann. Math. Statist. 41, 369-375.

[5] Paik, U. B. and Federer, W. T. [1970], "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.

[6] Raktoe, B. L. and Federer, W. T. [1970], "On the number of generators of saturated main effect fractional replicates." (in press)

[7] Zacks, S. [1964], "General least squares estimators for randomized fractional replication designs." Ann. Math. Statist. 35, 696-704.