RANDOMIZATION AND FACTORIAL EXPERIMENTS

By S. Ehrenfeld and S. Zacks

New York University and Technion, Israel Institute of Technology

1. Introduction and Summary. Many problems in experimental design can be stated as follows: An experimenter can perform N trials to estimate v parameters $\beta_1, \beta_2, \cdots, \beta_v$. There is available a set, X, of treatment combinations which may be performed (allowing repetitions). At each trial a treatment combination, x, is chosen from the set X and applied to an experimental unit. Thus, for each treatment combination there is associated a random variable $Y(x)$ whose distribution may depend on parameters $\beta_1, \cdots, \beta_k (k \geqq v)$ and on x. That is, Prob $(Y(x) \leqq t) = F(t \mid \beta, x)$ where $\beta' = (\beta_1, \beta_2, \cdots, \beta_k)$. The problem is how to choose treatment combinations $x_1, x_2, \cdots, x_N$ in the set X, allowing repetitions, to observe $Y(x_1), \cdots, Y(x_N)$ and to make inferences concerning the β's.

In this paper we consider a special but important case. It is usually assumed that the number of available experiments, N, is larger than the number of parameters, i.e., $N > k$. For factorial experiments this is often not the case and N may be substantially less than k but still larger than v, the number of β's of particular interest. In a sense the β's not of interest are nuisance parameters. For example, in 2^m factorial experiments, the set X consists of $k = 2^m$ factorial combinations. The random variable, $Y(x)$, associated with each of these k combinations, depends on k parameters, one for the mean, and the other for the $k - 1$ orthogonal contrasts corresponding to the main effects and various interactions.

The "classical" approach to the case $N < k$ is through the fractional factorial designs, where the parameters of interest are confounded with effects assumed negligible, see [3], [5], [6], [8], [15]. These designs are often used for exploratory purposes, where one wishes to consider many possible factors, and where interactions, even of high order, cannot always be assumed negligible.

In this paper we study two randomization procedures for p^m factorial experiments where one obtains unbiased estimates, valid tests and confidence intervals for parameters of interest without the usual assumptions concerning interactions. These designs, called *Randomized Fractional Factorials*, consist of choosing $x_1, x_2, \cdots, x_N$ in X in some randomized manner. Randomization plays a vital part in modern statistics. Early work in this connection is by R. A. Fisher [7]. More detailed discussions are given by E. J. G. Pitman [10], M. B. Wilk and O. Kempthorne [13], J. Cornfield and J. W. Tukey [4] and others [9], [12], [13], [14]. Much of the work in this area concerns randomization with respect to the experimental units in the experiment. Recently, increased consideration has

Received February 2, 1960; revised July 5, 1960.

been given to randomization with regard to the choice of treatment combinations. The designs developed, called *random balance designs* have been studied by Satterthwaite [11]. A critical discussion of various aspects of this work is given in [11]. The motivation of this work is mostly that of screening the interesting parameters from all the possible ones. In the present paper, the motivation is somewhat different and is concerned with inference about certain pre-assigned parameters out of a total of p^m parameters. Let us suppose that the experimenter is particularly interested in p^s out of p^m, $(m > s)$ parameters. Two randomization procedures are studied in detail. Randomization Procedure I, discussed in Section 3, is to choose at random, with or without replacement, blocks of treatment combinations, out of p^{m-s} ones constructed by confounding the "nuisance" parameters. In other words, the set X is divided into p^{m-s} blocks, according to the usual fractional factorial schemes and some of the sets are chosen at random. Such a procedure is suggested by Cochran and Cox in [3]. Randomization Procedure II, discussed in Section 4, is to choose at random treatments from every block. In this case, however, the blocks are constructed by confounding the p^s chosen parameters, and not the "nuisance" ones. It is proved that in randomized fractional designs from a 2^m system, the second procedure gives estimates of all the chosen parameters with equal variance, while the first may estimate different parameters with different variances. In the case p^m $(p \geqq 3)$ both procedures may estimate with unequal variances. In both procedures, however, with some replication, still keeping the total number of experiments $< p^m$ one can test hypotheses and obtain confidence intervals for the chosen parameters. Analysis of variance tables are derived and various tests of hypotheses, suggested by the usual F-like ratios, are indicated. The properties of these tests and distribution problems will be studied in a subsequent paper. The analysis of variance also provides a method for testing whether the p^{m-s} parameters not chosen are significantly different from zero.

To illustrate the procedures consider the simple case of four factors, A, B, C, and D with $p = s = 2$. That is, there are four parameters of interest. Let these be the mean M, ABC, CD and ABD. These four are considered to emphasize that they are quite arbitrary except for the requirement that they be a group under the usual multiplication rule.

In method I we divide the sixteen treatment (2^4) combinations into four sets (2^{4-2}) each composed of four combinations (2^2) according to a defining relationship which holds between certain "nuisance" parameters. There are several such possible relationships (in this case 66) but we will choose one for illustration. With the usual notation we can have, using A, B, and AB in the defining relationship,

$$\begin{aligned} I &= A = B = AB \\ I &= A = -B = -AB \\ I &= -A = B = -AB \\ I &= -A = -B = AB. \end{aligned}$$

The treatment combinations in the sets are

(1)	ab	a	b
c	abc	ac	bc
d	abd	ad	bd
cd	$abcd$	acd	bcd.

We choose two sets at random, either with or without replacement, and combine the estimates from each set. The estimate for ABC, $\widehat{ABC}$, for example, is unbiased with variance

$$V(\widehat{ABC}) = \sigma^2/8 + w[(BC)^2 + (AC)^2 + (C)^2],$$

where w is equal to $\frac{1}{2}$ or $\frac{1}{3}$ according to whether the sampling is with or without replacement.

An analysis of variance for testing the hypothesis $ABC = 0$ can be obtained by comparing the mean square for ABC with the mean square associated with the variation between the estimates in the two chosen sets. Similar remarks hold for estimating all the parameters of interest. Note that, in Procedure I, the division of the treatment combinations into sets depends on the defining relationship.

In method II we divide the treatments into four sets (2^2) of four (2^{4-2}) using the parameters of interest in the defining relationship

$$\begin{aligned} I &= ABC = CD = ABD \\ I &= ABC = -CD = -ABD \\ I &= -ABC = CD = -ABD \\ I &= -ABC = -CD = ABD. \end{aligned}$$

These sets are

(1)	abc	cd	abd
ab	c	$abcd$	d
acd	bd	a	bc
bcd	ad	b	ac.

In this case, however, we choose a random sample, say two treatment combinations, at random, from each set. The estimates are obtained by taking the appropriate contrasts of the observation totals for each set. The estimates are unbiased with constant variance V, where

$$\begin{aligned} V = \sigma^2/8 + v[&(A)^2 + (B)^2 + (AB)^2 + (C)^2 + (AC)^2 + (BC)^2 \\ &+ (D)^2 + (AD)^2 + (BD)^2 + (ACD)^2 + (BCD)^2 + (ABCD)^2], \end{aligned}$$

and where v is equal to $\frac{1}{8}$ or $\frac{1}{12}$ according to whether sampling is with or without replacement. An analysis of variance scheme can be used by comparing the mean square of the estimate with the mean square associated with the variation between choices within sets.

In the above, we have a $\frac{1}{2}$ fractional factorial. A $\frac{3}{4}$ fractional factorial is obtained by choosing three blocks in method I, or three combinations per block in method II. In the subsequent development we also consider replications at each chosen treatment combination. This leads to a test of the hypothesis concerning the significance of the nuisance parameters.

It is interesting to note, in this example, that method II gives a constant variance while method I does not. For any possible choice of defining relationship, if the variances of the estimators of the interesting parameters β_i $(i = 1, \cdots, 4)$ are denoted by $V_I(\hat{\beta}_i)$ for method I, then the constant V of method II is equal to the average $\sum V_I(\hat{\beta}_i)/4$. This result is generally true for $p = 2$, but fails to hold when $p > 2$.

Randomization Procedure I is essentially a cluster type of sampling from the population of treatment combinations X. Randomization Procedure II is essentially a stratified type of sampling. The population X is divided into p^s strata and from each stratum a random sample is drawn. It should be emphasized that for both procedures, the sub-division of X can be obtained by the usual, standard confounding methods.

The methods developed in this paper can easily be generalized to the case of mixed factorial experiments. In Section 5 a discussion of various questions raised in this paper is given. Some of these are concerned with confidence intervals, distribution problems, the comparison of Procedures I and II, and with a comparison between randomized and non-randomized designs.

Finally the procedures developed lend themselves to a sequential approach where at each stage a decision is made about the importance of the "nuisance" parameters. Furthermore, the sequence of steps can be carried out keeping the necessary "orthogonality" properties.

2. Basic Notions and the Statistical Model. A p^m factorial system is a system comprised of m factors, each at p levels. It will be assumed that p is a prime number. The space of treatment combinations, X, is represented by the set $X = ((i_0, i_1, \cdots, i_{m-1}) : i_j = 0, 1, \cdots, p-1 \text{ for all } j = 0, \cdots, m-1)$ which, clearly, contains p^m points. The jth coordinate of a point represents the i_jth level of factor j. A *standard order* of the points x in X, is given by the relationship between the coordinates of a point $x_v \equiv (i_0, i_1, \cdots, i_{m-1})$ and the order subscript

$$v = \sum_{j=0}^{m-1} i_j p^j.$$

This order relationship between the points of the treatment space is unique. It is similar to that given by F. Yates in his procedure [15].

The multiplication operator $\otimes$ between any two treatment combinations x and x' is defined as follows: If $x \equiv (i_0, i_1, \cdots, i_{m-1})$ and $x' \equiv (i_0', i_1', \cdots, i_{m-1}')$ then $x \otimes x' \equiv (i_0'', i_1'', \cdots, i_{m-1}'')$ where, $i_j'' \equiv i_j + i_j' \pmod{p}$ for all $j = 0, \cdots, m-1$. It follows, immediately, that the set X is a group with respect to the operator $\otimes$.

The order of a treatment combination x_t obtained by multiplying x_v by x_u is given as follows: If $v = \sum_{j=0}^{n-1} i_j p^j$ and $u = \sum_{j=0}^{m-1} i'_j p^j$, then $t = \sum_{j=0}^{m-1} k_j p^j$, where $k_j \equiv i_j + i'_j \pmod{p}$. We designate this relationship between v, u and t by: $t = u \oplus v$. We denote by $[x_v]^a$ $(a = 0, \cdots, p-1)$, the multiplication of x_v by itself a-times. Also, $[x_v]^0 = x_o$ where $x_o \equiv (0, \cdots, 0)$.

A treatment x_u is said to be *independent* of a set of treatments $x_{v_1}, x_{v_2}, \cdots, x_{v_n}$ if there are no n numbers $a_1, a_2, \cdots, a_n$ such that,

$$x_u = [x_{v_1}]^{a_1} \otimes [x_{v_2}]^{a_2} \otimes \cdots \otimes [x_{v_n}]^{a_n}.$$

Every group of p^k treatments is generated by k independent treatments. We now specify the statistical model for the p^m factorial system. Let $Y(x_v)$ be a random variable associated with the treatment combination x_v which measures the response of the system to treatment combination x_v. The relationship between the expected value of each random variable $Y(x_v)$ and treatment x_v is given by a linear function of parameters $\beta_0, \beta_1, \cdots, \beta_{p^m-1}$ as follows:

$$(2.1) \qquad E(Y(x_v)) = \sum_{u=0}^{p^m-1} c_u(x_v)\beta_u \quad \text{for every} \quad v = 0, \cdots, p^m - 1.$$

The parameters β_u have the usual interpretation of main effects and interactions of the m factors. We distinguish between linear effects, quadratic effects and effects of higher order. We also distinguish between linear-linear interactions, linear-quadratic, etc. A discussion of this model is given in [8]. We further describe the structure of the p^m parameters, β_u, by considering the space B of p^m points where,

$$B = \{(\lambda_0, \lambda_1, \cdots, \lambda_{m-1}) : \lambda_j = 0, \cdots, p-1 \quad \text{for all} \quad j = 0, \cdots, m-1\}.$$

The correspondence between the parameters β_u and the points of B is given by the usual standard order relation specified by, $u = \sum \lambda_j p^j$.

We introduce the multiplicative operator $\otimes$ on the space B. The unit element of this group $\beta_0 \equiv (0, 0, \cdots, 0)$ is the mean response of all the treatment combinations. The parameters $\beta_{p^k} \equiv (0, \cdots, 1, 0, \cdots, 0)$, $(k = 0, \cdots, m-1)$, where the one is in the kth place, corresponds to the main effects. Linear interactions correspond to points where coordinates are zero or one with at least two coordinates ones.

According to the usual interpretation of the β's it can be shown that the coefficients $c_u(x_v)$ of the linear system (2.1) are related to the coefficients of the orthogonal polynomials of order p, by the following relation: Let us denote by $\mathbf{C}^{(p^m)}$ the matrix of coefficients $c_u(x_v)$ of system (2.1). Furthermore, let $\mathbf{C}^{(p)}$ be the matrix whose column vectors are the coefficients of orthogonal polynomials of order p-namely,

$$\mathbf{C}^{(p)} = \begin{bmatrix} 1 & \xi_{0,1} & \cdots & \xi_{0,p-1} \\ 1 & \xi_{1,1} & \cdots & \xi_{1,p-1} \\ \vdots & \vdots & & \vdots \\ 1 & \xi_{p-1,1} & \cdots & \xi_{p-1,p-1} \end{bmatrix}.$$

The inner product of any two different column vectors of $\mathbf{C}^{(p)}$ is zero. The matrix $\mathbf{C}^{(p^m)}$ can be defined recursively for all $m \geqq 2$ by,

$$\mathbf{C}^{(p^m)} = \begin{bmatrix} \mathbf{C}^{(p^{m-1})} & \xi_{0,1}\mathbf{C}^{(p^{m-1})} & \cdots & \xi_{0,p-1}\mathbf{C}^{(p^{m-1})} \\ \mathbf{C}^{(p^{m-1})} & \xi_{1,1}\mathbf{C}^{(p^{m-1})} & \cdots & \xi_{1,p-1}\mathbf{C}^{(p^{m-1})} \\ \vdots & \vdots & & \\ \mathbf{C}^{(p^{m-1})} & \xi_{p-1,1}\mathbf{C}^{(p^{m-1})} & \cdots & \xi_{p-1,p-1}\mathbf{C}^{(p^{m-1})} \end{bmatrix}. \tag{2.2}$$

In other words, the matrix $\mathbf{C}^{(p^m)}$ is obtained from $\mathbf{C}^{(p^{m-1})}$ by a Kronecker's direct multiplication of $\mathbf{C}^{(p^{m-1})}$ by $\mathbf{C}^{(p)}$ from the left, i.e., $\mathbf{C}^{(p^m)} = \mathbf{C}^{(p)} \otimes \mathbf{C}^{(p^{m-1})}$. For example,

(i) when $p = 2$, $\mathbf{C}^{(2)} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, and thus

$$\mathbf{C}^{(2^m)} = \begin{bmatrix} \mathbf{C}^{(2^{m-1})} & -\mathbf{C}^{(2^{m-1})} \\ \mathbf{C}^{(2^{m-1})} & \mathbf{C}^{(2^{m-1})} \end{bmatrix}; \tag{2.3}$$

(ii) when $p = 3$, $\mathbf{C}^{(3)} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}$, and thus

$$\mathbf{C}^{(3^m)} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix} \otimes \mathbf{C}^{(3^{m-1})}. \tag{2.4}$$

In order to simplify the further development in treating randomization procedures it is necessary to study the structure of the matrices $\mathbf{C}^{(p^m)}$. We now derive some properties of these matrices.

From relationship (2.2) and the properties of $\mathbf{C}^{(p)}$ we easily obtain that the column vectors of $\mathbf{C}^{(p^m)}$ are orthogonal. Thus, $(\mathbf{C}^{(p^m)})'\ (\mathbf{C}^{(p^m)}) = \mathbf{\Delta}^{(p^m)}$, where $\mathbf{\Delta}^{(p^m)}$ is a non-singular diagonal matrix. Moreover, as a direct consequence of the recursion relation and the associative property of the Kronecker direct multiplication operator, $\otimes$, we have that, for every $1 \leqq s \leqq m$, the relationship between $C^{(p^s)}$ and $C^{(p^m)}$ is given by $\mathbf{C}^{(p^m)} = \mathbf{C}^{(p^{m-s})} \otimes \mathbf{C}^{(p^s)}$.

LEMMA. *The elements of the matrix* $\mathbf{C}^{(p^m)}$ *are related to those of* $\mathbf{C}^{(p^s)}$ *and* $\mathbf{C}^{(p^{m-s})}$ *according to the following*:

$$c^{(p^m)}_{i+jp^s,l} = c^{(p^{m-s})}_{j,q_l} \cdot c^{(p^s)}_{i,r_l} \tag{2.5}$$

for all $j = 0, \cdots, p^{m-s} - 1$, *and* $i = 0, \cdots, p^s - 1$, *where*

$$l = q_l p^s + r_l \qquad (q_l = 0, \cdots, p^{m-s} - 1;\ r_l = 0, \cdots, p^s - 1).$$

PROOF: We have that $\mathbf{C}^{(p^m)} = \mathbf{C}^{(p^{m-s})} \otimes \mathbf{C}^{(p^s)}$. By this structure the matrix $\mathbf{C}^{(p^m)}$ is divided into $p^{m-s} \times p^{m-s}$ submatrices, given by $c^{(p^{m-s})}_{j,q_l}\mathbf{C}^{(p^s)}$.

Thus, the element of $\mathbf{C}^{(p^m)}$ in the vth and lth column belongs to the submatrix $c^{(p^{m-s})}_{j,q_l}\mathbf{C}^{(p^s)}$, where $j = [v/p^s]$ and $q_l = [l/p^s]$, where $[x]$ means the largest integer

not greater than x. Moreover, if $v = i + jp^s$ $(i = 0, \cdots, p^s - 1)$, and $l = r_l + q_l p^s$ $(r_l = 0, \cdots, p^s - 1)$, then $c_{vl}^{(p^m)}$ is the element in the ith row and r_lth column of $c_{i,q_l}^{(p^{m-s})}\mathbf{C}^{(p^s)}$.

By the lemma proved here, the following well known relationship can be easily shown, namely: In a 2^m factorial system, the elements of the matrix $\mathbf{C}^{(2^m)}$ in any row, and in columns corresponding to β_{l_1}, β_{l_2} and β_k, where $\beta_k = \beta_{l_1} \otimes \beta_{l_2}$ $(l_1, l_2, = 0, \cdots, p^m - 1)$ are related by the rule $c_k^{(2^m)}(x) = c_{l_1}^{(2^m)}(x)c_{l_2}^{(2^m)}(x)$. It should be emphasized that this, being true for 2^m factorial system, is not necessarily true for the general case of p^m factorial systems, where $p \geqq 3$.

LEMMA: *In a 2^m factorial system, the value of the coefficients*

$$c_u^{(2^m)}(x_v) \quad \textit{for} \quad v = \sum i_j 2^j \quad \textit{and} \quad u = \sum \lambda_j 2^j \qquad (i_j, \lambda_j = 0, 1)$$

is given by:

$$c_u^{(2^m)}(x_v) = (-1)^{\sum_{j=0}^{m-1} \lambda_j (1 - i_j)}. \tag{2.6}$$

PROOF: Every parameter β_u $(u = 0, 1, \cdots, 2^m - 1)$ can be represented as

$$\beta_u = [\beta_1]^{\lambda_0} \otimes [\beta_2]^{\lambda_1} \otimes \cdots \otimes [\beta_{2^{m-1}}]^{\lambda_{m-1}}, \tag{2.7}$$

where β_{2^k} $(k = 0, 1, \cdots, m - 1)$ are independent parameters that generate the parameter group (the "main effect" parameters). From relation (2.7) and the preceding remarks, it follows that for every treatment combination x_v,

$$c_u^{(2^m)}(x_v) = [c_1^{(2^m)}(x_v)]^{\lambda_0}[c_2^{(2^m)}(x_v)]^{\lambda_1} \cdots [c_{2^{m-1}}^{(2^m)}(x_v)]^{\lambda_{m-1}}. \tag{2.8}$$

The matrix $\mathbf{C}^{(2^2)}$ reveals that if $x_v \equiv (i_0, i_1)$ then,

$$c_{2^k}^{(2^2)}(x_v) = (-1)^{(1-i_k)} \quad \text{for} \quad k = 0, 1. \tag{2.9}$$

Let us prove, by induction, that (2.9) is true for all m.

Assuming that $c_{2^k}^{(2^{m-1})}(x_v) = (-1)^{(1-i_k)}$ when $k = 0, \cdots, m-2$, and applying relation (2.3) since $\mathbf{C}^{(2^m)} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \otimes \mathbf{C}^{(2^{m-1})}$, we obtain for the parameter $\beta_{2^{m-1}}$,

$$c_{2^{m-1}}^{(2^m)}(x_v) = \begin{cases} -1 & \text{if} \quad v = 0, 1, \cdots, 2^{m-1} - 1 \\ 1 & \text{if} \quad v = 2^{m-1}, 2^{m-1} + 1, \cdots, 2^m - 1. \end{cases} \tag{2.10}$$

For the other independent parameters β_k $(k = 0, 1, \cdots, m - 2)$ the following relation holds:

$$c_{2^k}^{(2^m)}(x_v) = \begin{cases} c_{2^k}^{(2^{m-1})}(x_v) & \text{if} \quad v = 0, 1, \cdots, 2^{m-1} - 1 \\ c_{2^k}^{(2^{m-1})}(x_{v-2^{m-1}}) & \text{if} \quad v = 2^{m-1}, \cdots, 2^m - 1, \end{cases} \tag{2.11}$$

moreover,

$$v \equiv \begin{cases} (i_0, i_1, \cdots, i_{m-2}, 0) & \text{if} \quad v < 2^{m-1} \\ (i_0, i_1, \cdots, i_{m-2}, 1) & \text{if} \quad v \geqq 2^{m-1}. \end{cases} \tag{2.12}$$

Hence, $c_{2^k}^{(2^m)}(x_v) = (-1)^{1-i_k}$ for all $k = 0, \cdots, m-1$. By substituting this result in (2.8), formula (2.6) is proved.

3. Randomization Procedure I in Fractional Replication Designs.

3.1. *The case of a $n/2^{m-s}$ replicate of a 2^m $(m > s)$ factorial design.* A $n/2^{m-s}$ replicate of a 2^m $(m > s)$ factorial system, according to Randomization Procedure I, is a design in which n blocks of treatment combinations are chosen, at random, out of 2^{m-s} blocks. The 2^{m-s} blocks, each containing 2^s treatment combinations, are constructed as follows:

(i) A subgroup of 2^s interesting parameters is chosen and specified.

(ii) A set of $(m-s)$ independent parameters is chosen. The parameters of this set do not belong to the chosen subgroup of interesting parameters. Designate these parameters by $(\beta_{d_0}, \beta_{d_1}, \cdots, \beta_{d_{m-s-1}})$.

(iii) Specify the subgroup of 2^{m-s} parameters generated by the basis $(\beta_{d_0}, \beta_{d_1}, \cdots, \beta_{d_{m-s-1}})$. Every parameter β_{d_u} $(u = 0, \cdots, 2^{m-s}-1)$ contained in this subgroup is obtained by multiplication of independent parameters. This subgroup is called the *defining subgroup.*

(iv) Classify all the treatment combinations into 2^{m-s} mutually exclusive blocks by the following rule: If $x \equiv (i_0, i_1, \cdots, i_{m-1})$ satisfies the following system of equations

$$\sum_{k=0}^{m-1} \lambda_{kd_j} i_k \equiv a_j \pmod 2 \quad \text{for all} \quad j = 0, 1, \cdots, m-s-1,$$

where $a_j = 0, 1$, then x belongs to X_v whose index v is given by

$$\sum_{j=0}^{m-s-1} a_j 2^j \qquad \text{where} \qquad \lambda_{kd_j} \ (k = 0, 1, \cdots, m-1)$$

are the coordinates of the independent defining parameters β_{d_j}, i.e., $\beta_{d_j} \equiv (\lambda_{0d_j}, \cdots, \lambda_{(m-1)d_j})$. This classification rule is common to procedures of confounded designs, see O. Kempthorne [8]. However, for the sake of further development of the theory the following definition of blocks of treatment combinations, X_v, is adopted.

$$\begin{aligned} X_v = \{x : c_{d_j}^{(2^m)}(x) &= (-1)^{(i_j - L(d_j))}, \quad j = 0, 1, \cdots, m-s-1, \\ \text{where } v &= \sum_{j=0}^{m-s-1} i_j 2^j, \qquad \beta_{d_j} \equiv (\lambda_{0d_j}, \lambda_{1d_j}, \cdots, \lambda_{(m-1)d_j}), \\ \text{and } L(d_j) &\equiv \sum_{k=0}^{m-1} \lambda_{kd_j} \pmod 2\}. \end{aligned} \tag{3.1}$$

Classification according to definition (3.1) is particularly convenient, since it can be carried out just by comparing the coefficients of $\mathbf{C}^{(2^m)}$ in different rows and columns corresponding to the independent defining parameters β_{d_j}. It can be easily shown that classification according to (3.1) and that given by solving the equations $\sum \lambda_{kd_j} i_k \equiv a_j \pmod 2$, $(j = 0, \cdots, m-s-1)$, are equivalent.

Let us classify all the parameters β_u into 2^{m-s} exclusive sets as follows:

(i) The first subset, B_0 will contain all the 2^s chosen parameters.

(ii) Let us order the chosen parameters, belonging to B_0 according to the order relationship prevailing by the standard order. Thus, the chosen parameters are, according to the order in B_0, $(\beta_0, \beta_{l_1}, \cdots, \beta_{l_{2^s-1}})$, where $l_r < l_{r+1}$ $(r = 0, \cdots, 2^s - 1)$.

(iii) Construct B_1 by multiplying all the chosen parameters β_{l_r} successively by the defining parameter β_{d_1}. All the parameters $\beta_{l_r} \otimes \beta_{d_1}$ $(r = 0, \cdots, 2^s - 1)$ constitute B_1.

The parameters obtained are called the *aliases* of β_{l_r} with respect to β_{d_1}. This relationship is denoted by $\beta_{1(l_r)} = \beta_{l_r} \otimes \beta_{d_1}$. All the subsequent sets are obtained similarly. Correspondingly, the aliases of β_{l_r} with respect to β_{d_u} are denoted by $\beta_{u(l_r)} = \beta_{l_r} \otimes \beta_{d_u}$.

Thus, all the 2^m treatment combinations are classified into 2^{m-s} blocks, X_v, and all the 2^m parameters are classified into 2^{m-s} subsets, B_u $(v, u = 0, 1, \cdots, 2^{m-s} - 1)$.

Since there is no restriction, whatsoever, on the choice of the group of interesting parameters, B_0, the development of a general theory requires that new matrices, denoted by $\mathbf{P}_{vu}^{(2^s)}$ be introduced $(v, u = 0, \cdots, 2^{m-s} - 1)$.

Thus, let us define the matrix $\mathbf{P}_{vu}^{(2^s)}$ to be a square matrix of order 2^s, whose elements are those of $\mathbf{C}^{(2^m)}$ corresponding to the treatments belonging to X_v and parameters belonging to B_u. The order of elements of $\mathbf{P}_{vu}^{(2^s)}$ is the same as that of its elements in $C^{(2^m)}$.

DEFINITION. An estimator of the 2^s chosen parameters, β, given the block X_v of treatments, and the independent defining parameters $(\beta_{d_0}, \cdots, \beta_{d_{m-s-1}})$ is as follows:

$$\hat{\beta}_{v(d)} = 2^{-s}(\mathbf{P}_{v0(d)}^{(2^s)})' y(X_v) \tag{3.2}$$

where $y(X_v)$ is the vector of random variables associated with the treatments in X_v and the subscript, in brackets, d, refers to the defining group. Different defining groups may, of course, lead to different estimates. In order to study the properties of this estimator the structure of the matrices $\mathbf{P}_{vu}^{(2^s)}$ must be examined.

THEOREM 3.1: *For every confounding system, given by a set of independent defining parameters* $(\beta_{d_0}, \beta_{d_1}, \cdots, \beta_{d_{m-s-1}})$ *the matrices*

$$\mathbf{P}_{vu}^{(2^s)} \quad (v, u = 0, 1, \cdots, 2^{m-s} - 1)$$

are related to $\mathbf{P}_{00}^{(2^s)}$ *by the following relationship*;

$$\mathbf{P}_{vu}^{(2^s)} = (-1)^{\sum_{j=0}^{m-s-1} i_j'(i_j - L(d_j))} \mathbf{P}_{00}^{(2^s)}, \tag{3.3}$$

where $v = \sum_{j=0}^{m-s-1} i_j 2^j$; $u = \sum_{j=0}^{m-s-1} i_j' 2^j$ *and* $L(d_j) \equiv \sum_{k=0}^{m-1} \lambda_{kd_j} \pmod 2$.

PROOF. Every defining parameter β_{d_u} $(u = 0, 1, \cdots, 2^{m-s} - 1)$ is given by

$$\beta_{d_u} = [\beta_{d_0}]^{i_0'} \otimes [\beta_{d_1}]^{i_1'} \otimes \cdots \otimes [\beta_{d_{m-s-1}}]^{i_{m-s-1}'}, \tag{3.4}$$

where $u = \sum i'_j 2^j$ and β_{d_j} $(j = 0, \cdots, m - s - 1)$ are independent. According to (3.1), the coefficients $c_{d_j}^{(2^m)}(x_v)$ associated with the independent defining parameters β_{d_j} $(j = 0, \cdots, m - s - 1)$ and with treatment combinations which belong to X_v are given by $c_{d_j}^{(2^m)}(x_v) = (-1)^{i_j - L(d_j)}$. Thus, from (3.3) and the recursion relation it follows that

$$c_{d_u}^{(2^m)}(x_v) = \prod_{j=0}^{m-s-1} [c_{d_j}^{(2^m)}(x_v)]^{i'_j} = (-1)^{\sum_{j=0}^{m-s-1} i'_j (i_j - L(d_j))} \tag{3.5}$$

However, the parameters that belong to subset B_u are related to those of B_0 by the relationship: $\beta_{u(l_r)} = \beta_{l_r} \otimes \beta_{d_u}$ where β_{l_r} in B_0. It follows that all the elements of $\mathbf{P}_{vu}^{(2^s)}$ are obtained by multiplying those of $\mathbf{P}_{00}^{(2^s)}$ by $c_{d_u}^{(2^m)}(x_v)$ correspondingly.

LEMMA: *The vectors of the matrices* $\mathbf{P}_{vu}^{(2^s)}$ $(v, u = 0, \cdots, 2^{m-s} - 1)$ *are orthogonal*, i.e.

$$(\mathbf{P}_{vu}^{(2^s)})'(\mathbf{P}_{vu}^{(2^s)}) = 2^s I^{(2^s)}, \tag{3.6}$$

where $I^{(2^s)}$ *is a unit matrix of order* 2^s. The proof of this lemma is straightforward.

Since the coefficients which relate the matrices $\mathbf{P}_{vu}^{(2^s)}$ to $\mathbf{P}_{00}^{(2^s)}$ according to (3.2) play an important role at the sequel, let us designate them by b_{vu}. Thus

$$b_{vu} = (-1)^{\sum_{j=0}^{m-s-1} i'_j (i_j - L(d_j))} \tag{3.7}$$

Clearly, b_{vu} depends on the coordinates of the independent defining parameter through $L(d_j)$. The matrix of the elements, b_{vu} will be designated by $B^{(2^{m-s})}(L(d_0), \cdots, L(d_{m-s-1}))$ to indicate this dependence. The above is a square matrix of order 2^{m-s}.

It is easily shown that every matrix $\mathbf{B}^{(2^{m-s})}(L(d_0), \cdots, L(d_{m+s+1}))$ is a permutation of the rows of $B^{(2^{m-s})}(0, 0, \cdots, 0)$. Also the vectors of $\mathbf{B}^{(2^{m-s})}(L(d_0), \cdots, L(d_{m-s-1}))$ are orthogonal for all the sets $(L(d_0), \cdots, L(d_{m-s-1}))$.

THEOREM 3.2: *For every given* X_v $(v = 0, 1, \cdots, 2^{m-s} - 1)$

$$E(\hat{\beta}_{v(d)} \mid X_v) = \beta + \sum_{u=1}^{2^{m-s}-1} b_{vu} \beta^*_{u(d)}, \tag{3.9}$$

where $\beta^*_{u(d)}$ *is a vector of the parameters, which belong to the subset* B_u.

PROOF. According to (3.2),

$$E(\hat{\beta}_{v(d)} \mid X_v) = 2^{-s}(\mathbf{P}_{v0(d)}^{(2^s)})' E(y(X_v)).$$

According to the linear model (2.1), $E(y(X_v)) = \sum_{u=0}^{2^{m-s}-1} (\mathbf{P}_{vu(d)}^{(2^s)}) \beta^*_{u(d)}$. Hence

$$E(\hat{\beta}_{v(d)} \mid X_v) = 2^{-s} \sum_u (\mathbf{P}_{v0(d)}^{(2^s)})'(\mathbf{P}_{vu(d)}^{(2^s)}) \beta^*_{u(d)}.$$

By Theorem 3.1 and (3.7) we have

$$E(\hat{\beta}_{v(d)} \mid X_v) = 2^{-s} \sum_u b_{v0} b_{vu} (\mathbf{P}_{00}^{(2^s)})'(\mathbf{P}_{00}^{(2^s)}) \beta^*_{u(d)} = \beta + b_{v0} \sum_u b_{vu} \beta^*_{u(d)}.$$

However, according to (3.7), $b_{v0} = 1$.

THEOREM 3.3: *If a block of treatment combinations is chosen at random then* $\hat{\beta}_{v(d)}$ *is unbiased and the variances of its components are*:

$$V(\hat{\beta}_{lv(d)}) = \sigma^2/2^s + \sum_u \beta^{*2}_{lu(d)}, \tag{3.10}$$

where $\hat{\beta}_{lv(d)}$ *is the* lth *component* ($l = 0, \cdots, 2^s - 1$) *of* $\hat{\beta}_{v(d)}$, *and* $\beta^*_{lu(d)}$ *is the* lth *component of* $\beta^*_{u(d)}$.

PROOF.

(i) The estimator $\hat{\beta}_{v(d)}$ is unbiased, because:

$$E(\hat{\beta}_{v(d)}) = E_v\{E(\hat{\beta}_{v(d)} \mid X_v)\} = \beta + \sum_u E_v(b_{vu})\beta^*_{u(d)}$$
$$= \beta + 2^{m-s} \sum_u (\sum_v b_{vu})\beta^*_{u(d)},$$

but $\sum_v b_{vu}$ is zero, as is seen by (3.7).

(ii) The variance of a parameter $\hat{\beta}_{lv(d)}$ is given by

$$V(\hat{\beta}_{lv(d)}) = E_v\{V(\hat{\beta}_{lv(d)} \mid X_v)\} + V_v\{E(\hat{\beta}_{lv(d)} \mid X_v)\}.$$

However,

$$E_v\{V(\hat{\beta}_{lv(d)} \mid X_v)\} = E_v\{\sigma^2/2^s\} = \sigma^2/2^s$$

and,

$$V_v\{E(\hat{\beta}_{lv(d)} \mid X_v)\} = V_v\{\beta + \sum_u b_{vu}\beta^*_{lu(d)}\}.$$

Since the matrix of b's is orthogonal, covariances of the b's are zero. Thus,

$$V_v\{E(\hat{\beta}_{lv(d)} \mid X_v)\} = \sum_{u=1}^{2^{m-s}-1} \beta^{*2}_{lu(d)} V_v(b_{vu}).$$

According to (3.7), $V_v(b_{vu}) = 1$. If n blocks are chosen at random, and the estimator $\hat{\beta}_{v(d)}$ is the arithmetic mean of the n individual estimators, then $V(\hat{\beta}_{v(d)}) = \sigma^2/n2^s + n^{-1}M \sum_u \beta^{*2}_{u(d)}$, where

$$M = \begin{cases} 1 & \text{if sampling is with replacement} \\ 1 - (n-1)/(2^{m-s} - 1) & \text{if sampling is without replacement.} \end{cases}$$

3.2 *The case of a* n/p^{m-s} *replicate of a* p^m *factorial experiment* ($m > s$; $p \geqq 3$). In the present section, Randomization Procedure I is applied to the p^m factorial system, when $p \geqq 3$. The derivation of the theory, for the present case, is faced with some complications which were not present in the case of 2^m factorial systems. When $p \geqq 3$ the matrix $C^{(p)}$ might contain some zero elements. Thus, if we are free to choose any subgroup of p^s parameters and classify the treatment combinations into p^{m-s} blocks, by confounding a subgroup of defining parameters, it might happen that the matrix of coefficients $\mathbf{P}^{(p^s)}_{v0}$ ($v = 0, \cdots, p^{m-s} - 1$) is singular.

EXAMPLE: Suppose $p = 3$, $m = 2$ and $s = 1$. Let us choose the following subgroup of parameters: $\beta_0 = M$, $\beta_4 = AB$ and $\beta_8 = A^2B^2$; and the following defining parameter $\beta_d = A$, where M, AB, A^2B^2 etc. are the usual notation for the parameters. That is, M denotes the mean, AB the linear interaction between

factors A and B, and A^2B^2 the interaction between the quadratic of A and the quadratic of B. Thus, the three blocks of treatment combinations are

$$X_0 = \{(i_0, i_1) : i_0 \equiv 0 \pmod 3\} = (x_0, x_3, x_6)$$

$$X_1 = \{(i_0, i_1) : i_0 \equiv 1 \pmod 3\} = (x_1, x_4, x_7)$$

$$X_2 = \{(i_0, i_1) : i_0 \equiv 2 \pmod 3\} = (x_2, x_5, x_8).$$

If we define an estimator like in (3.2) we are faced with the problem that $\mathbf{P}_{10}^{(3)}$ is a singular matrix.

In order to avoid possible singularities, let us rule that every chosen subgroup of parameters should be generated by s main effects, i.e., $(\beta_{pi_0}, \beta_{pi_1}, \cdots, \beta_{pi_{s-1}})$ where $(j = 0, \cdots, m - 1)$.

For this class of subgroups, of order p^s, there is no loss in generality if we assume that the chosen sub-group of parameters is the set of first p^s parameters $(\beta_0, \beta_1, \cdots, \beta_{p^s-1})$ because here it is a matter of relabelling the parameters' order and those of the treatment combinations correspondingly in order to get a matrix of coefficients identical with $\mathbf{C}^{(p^m)}$.

Substantial difficulties may also arise, in the case $p \geqq 3$, if we choose the defining parameters without any restrictions. For example, take the case of $p = 3$, $m = 2$, $s = 1$. Let the chosen subgroup of parameters be (M, A, A^2) and let the defining parameter be $\beta_d = AB$. In this case it can be shown that the submatrices $\mathbf{P}_{v1}^{(3)}$ $(v = 0, 1, 2)$ have nonorthogonal column vectors.

In order to avoid complications of this kind, let us rule that the defining group should be generated by the $(m - s)$ independent parameters, representing main effects, which are not in the chosen group of parameters. Without loss of generality, let us assume that the defining group is generated by the set

$$(\beta_{p^s}, \beta_{p^s+1}, \cdots, \beta_{p^m-1}).$$

DEFINITION: An estimator of the vector of the chosen parameters β_0 given a block of treatment combinations X_v $(v = 0, \cdots, p^{m-s} - 1)$, is defined by

$$\hat{\beta}_{v_0} = (\mathbf{\Delta}^{(p^s)})^{-1}(\mathbf{C}^{(p^s)})' y(X_v). \tag{3.11}$$

THEOREM 3.4: *For any given block of treatment combinations*

$$X_v \ (v = 0, 1, \cdots, p^{m-s} - 1),$$

the conditional expection of $\hat{\beta}_{v_0}$ *is*

$$E(\hat{\beta}_{v_0} \mid X_v) = \beta_0 + \sum_{u=1}^{p^{m-s}-1} c_{vu}^{(p^{m-s})} \beta_u^*, \tag{3.12}$$

where β_u^* *are vectors of parameters alias to those of* β_0 *with respect to* β_{up^s}.

PROOF. According to the linear model (2.1),

$$E(y(X_v)) = \sum_{u=0}^{p^{m-s}-1} c_{vu}^{(p^{m-s})} (\mathbf{C}^{(p^s)}) \beta_u^*.$$

Substituting $E(y(X_v))$ into (3.11), we obtain

$$E(\hat{\beta}_{v_0} \mid X_v) = (\Delta^{(p^s)})^{-1} \sum_{u=0}^{p^{m-s}-1} c_{vu}^{(p^{m-s})} (\mathbf{C}^{(p^s)})'(\mathbf{C}^{(p^s)})\beta_u^* = \beta_0 + \sum_u c_{vu}^{(p^{m-s})}\beta_u^* .$$

If n blocks of treatment combinations, $X_{v_1}, X_{v_2}, \cdots, X_{v_n}$ are chosen, an estimator of β_0 is given by

$$\hat{\beta}_0 = \sum_{j=1}^{n} \hat{\beta}_{v_j 0}/n. \tag{3.13}$$

THEOREM 3.5: *If n blocks of treatment combinations are chosen at random, then $\hat{\beta}_0$ is an unbiased estimator of β_0 with a variance of its lth component ($l = 0, \cdots, p^s - 1$) given by*

$$V(\hat{\beta}_{l_0}) = \sigma^2/nd_l^{(p^s)} + M \sum_{u=1}^{p^{m-s}-1} d_u^{(p^{m-s})}\beta_{l+up^s}^2/np^{m-s}, \tag{3.14}$$

where

$$M = \begin{cases} 1, & \text{if sampling is with replacement} \\ 1 - (n-1)/(p^{m-s}-1), & \text{if sampling is without replacement.} \end{cases}$$

PROOF.

(i) The estimator $\hat{\beta}_0$ is unbiased, since

$$E(\hat{\beta}_0) = E_v\{E(\hat{\beta}_{v0} \mid X_v)\} = E_v\left\{\beta_0 + \sum_{u=1}^{p^{m-s}-1} c_{vu}^{(p^{m-s})}\beta_u^*\right\}$$

$$= \beta_0 + \sum_{u=1}^{p^{m-s}-1} \beta_u^* E_v(c_{vu}^{(p^{m-s})})$$

$$= \beta_0 + \frac{1}{p^{m-s}} \sum_{u=1}^{p^{m-s}-1} \beta_u^* \left\{\sum_{v=0}^{p^{m-s}-1} c_{vu}^{(p^{m-s})}\right\}.$$

Since $\sum_{v=0}^{p^{m-s}-1} c_{vu}^{(p^{m-s})} = 0$ for all $u = 1, 2, \cdots, p^{m-s} - 1$, we obtain $E(\hat{\beta}_{v0}) = \beta_0$. Moreover, from the unbiasedness of every $\hat{\beta}_{v_0}$ it immediately follows that $\hat{\beta}_0$ is unbiased.

(ii) The variance of the lth component of the vector β_0 is found as follows:

$$V\{\hat{\beta}_{l_0}\} = E_v\{V(\hat{\beta}_{l_0} \mid X_{v_1}, X_{v_2}, \cdots, X_{v_n})\} + V_v\{E(\hat{\beta}_{l_0} \mid X_{v_1}, \cdots, X_{v_n})\}.$$

The statistical model states that $y(X_v) = E(y(X_v)) + e$. Thus, according to (3.11) $\hat{\beta}_{v_0} = (\Delta^{(p^s)})^{-1}(\mathbf{C}^{(p^s)})'E(y(X_v)) + (\Delta^{(p^s)})^{-1}(\mathbf{C}^{(p^s)})'e$. Hence, the conditional variance of $\hat{\beta}_{l_0}$, given a block X_{v_j}, is

$$V(\hat{\beta}_{l_0} \mid X_{v_j}) = \sigma^2/d_l^{(p^s)} \quad \text{for all} \quad j = 1, 2, \cdots, n.$$

According to (3.1), $V(\hat{\beta}_{l_0} \mid X_{v_1}, \cdots, X_{v_n}) = \sigma^2/nd_l^{(p^s)}$. Furthermore, by Theorem 3.4

$$E\{\hat{\beta}_{l_0} \mid X_{v_1}, \cdots, X_{v_n}\} = \beta_{l_0} + n^{-1} \sum_{j=1}^{n} \sum_{u=1}^{p^{m-s}-1} c_{v_j u}^{(p^{m-s})}\beta_{l+up^s} .$$

Thus, from the orthogonality of the column vectors of $\mathbf{C}^{(p^{m-s})}$, we obtain:

$$V_v\{E(\hat{\beta}_{l_0} \mid X_{v_1}, \cdots, X_{v_n})\} = \sum_{u=1}^{p^{m-s}-1} \beta_{l+up^s}^2 V_v\left\{n^{-1}\sum_{j=1}^{n} c_{v_j u}^{(p^{m-s})}\right\}.$$

But,

$$V_v\left\{n^{-1}\sum_{j=1}^{n} c_{v_j u}^{(p^{m-s})}\right\} = \left(\frac{M}{np^{m-s}}\right)\sum_{v=0}^{p^{m-s}-1}(c_{vu}^{(p^{m-s})})^2 = \left(\frac{M}{np^{m-s}}\right) d_u^{(p^{m-s})},$$

where M is defined above.

3.3 *Testing of Hypotheses according to Randomization Procedure* I. In this section, we study procedures for testing the following types of hypotheses:

(i) $H_0: \beta_l = 0$ $(l = 0, \cdots, p^s - 1)$ against the alternative $H_1: \beta_l \neq 0$.

(ii) $H_{0u(l)}^*: \beta_{u(l)} = 0$ for *all* $u = 1, \cdots, p^{m-s} - 1$ against the alternative, $H_{1u(l)}^*$: at least one $\beta_{u(l)} \neq 0$ $(u = 1, \cdots, p^{m-s} - 1)$.

These form a set of hypotheses where $l = 0, \cdots, p^s - 1$.

Test statistics, for testing the above hypotheses, are suggested by an analysis of variance scheme, in which the sum of squares of deviations of all the random variables about the grand mean is partitioned into components, according to different sources of variation. In order to be able to test hypotheses of type (i) and (ii) we require that the number of chosen blocks $n \geqq 2$; and the number of repetitions of every chosen treatment combination $r \geqq 2$. If $n \geqq 2$ and $r = 1$ hypothesis of type (i) can still be tested.

The sum of squares, of deviations, of all the y's is partitioned as usual, into the sum of squares "within treatments" and "between treatments". The estimators $\hat{\beta}_{lv_j}$ $(l = 0, \cdots, p^s - 1)$ are the orthogonal contrasts between the treatments means of block X_{v_j} $(j = 1, \cdots, n)$. Every contrast of this kind carries one degree of freedom. As defined in (3.13), $\hat{\beta}_{l\cdot}$ is the mean of all $\hat{\beta}_{lv_j}$ over all the n chosen blocks. Thus, the quadratic forms

$$Q(\hat{\beta}_{l\cdot}) = rd_l^{(p^s)}\sum_{j=1}^{n}(\hat{\beta}_{lv_j} - \hat{\beta}_{l\cdot})^2 \qquad (l = 0, \cdots, p^s - 1) \tag{3.15}$$

carry $(n - 1)$ degrees of freedom.

It is obvious that all $Q(\hat{\beta}_{l\cdot})$ are mutually orthogonal. Clearly, $Q(\hat{\beta}_0)$ measures the variability between the defining parameters of the chosen blocks. $Q(\hat{\beta}_1)$ measures the variability between the aliases to β_1 in the chosen blocks, etc. Let us define, for all $l = 0, \cdots, p^s - 1$,

$$Q^*(\hat{\beta}_{l\cdot}) = rnd_l^{(p^s)}\hat{\beta}_{l\cdot}^2 . \tag{3.16}$$

Thus the F-like ratio

$$F_l^* = (n - 1)Q^*(\hat{\beta}_{l\cdot})/Q(\hat{\beta}_{l\cdot}), \quad (l = 0, \cdots, p^s - 1) \tag{3.17}$$

could serve as a test statistic, for $H_0: \beta_l = 0$ against $H_1: \beta_l \neq 0$. It is also seen that the test statistic

$$F_l = Q(\hat{\beta}_{l\cdot})/s_w^2(n - 1) \qquad (l = 0, \cdots, p^s - 1), \tag{3.18}$$

where $np^s(r-1)s_w^2 = \sum_{l=0}^{p^s-1}\sum_{j=1}^{n}\sum_{h=1}^{r}(y_{lv_jh} - y_{lv_j\cdot})^2$ could serve to test the hypothesis:

$H_{0u(l)}^*$: all the alias parameters, $\beta_{u(l)}$, are equal to zero; against the alternative:

$H_{1u(l)}^*$: at least one alias parameter $\beta_{u(l)}$ $(u = 1, \cdots, p^{m-s}-1)$ is not zero.

According to the theory of simple random sampling, if the sampling of blocks, X_{v_j}, is at random, with replacement then

$$(3.19)\quad E\{Q(\hat{\beta}_{l\cdot})/(n-1)\} = \sigma^2 + (rd_l^{(p^s)}/p^{m-s})\sum_{u=1}^{p^{m-s}-1} d_u^{(p^{m-s})}\beta_{l+up^s}^2$$

for all $l = 0, \cdots, p^s - 1$. It is easily shown that $E(s_w^2) = \sigma^2$. The results lead to the conclusion that the test statistic F_l could test the hypothesis $H_{0u(l)}^*$ against $H_{1u(l)}^*$. From Theorem (3.5) it follows that,

$$(3.20)\quad E\{rnd_l^{(p^s)}\hat{\beta}_{l\cdot}\} = \sigma^2 + (rd_l^{(p^s)}/p^{m-s})\sum_{u=1}^{p^{m-s}-1} d_u^{(p^{m-s})}\beta_{l+up^s}^2 + rnd_l^{(p^s)}\beta_l^2 .$$

TABLE 1

Analysis of Variance for Randomization Procedure I

Source of Variation	d.f.	S.S.	E(M.S.)
β_1	1	$Q^*(\hat{\beta}_{1\cdot})$	$\sigma^2 + (r\,d_1^{(p^s)}/p^{m-s})\sum_u d_u^{(p^{m-s})}\beta_{up^s+1}^2 + rn\,d_1^{(p^s)}\beta_1^2$
⋮	⋮	⋮	⋮
β_{p^s-1}	1	$Q^*(\hat{\beta}_{p^s-1\cdot})$	$\sigma^2 + (r\,d_{p^s-1}^{(p^s)}/p^{m-s})\sum_u d_u^{(p^{m-s})}\beta_{up^s-1}^2 + rn\,d_{p^s-1}^{(p^s)}\beta_{p^s-1}^2$
defining parameters	$n-1$	$Q(\hat{\beta}_{0\cdot})$	$\sigma^2 + (r/p^{m-2s})\sum_u d_u^{(p^{m-s})}\beta_{up^s}^2$
aliases to β_1	$n-1$	$Q(\hat{\beta}_{1\cdot})$	$\sigma^2 + (r\,d_1^{(p^s)}/p^{m-s})\sum_u d_u^{(p^{m-s})}\beta_{up^s+1}^2$
⋮	⋮	⋮	⋮
aliases to β_{p^s-1}	$n-1$	$Q(\hat{\beta}_{p^s-1\cdot})$	$\sigma^2 + (r\,d_{p^s-1}^{(p^s)}/p^{m-s})\sum_u d_u^{(p^{m-s})}\beta_{(u+1)p^s-1}^2$
all the chosen parameters	np^s-1	—	—
between treatments	np^s-1	$r\sum_{l=0}^{p^s-1}\sum_{j=1}^{n}(y_{lv_j\cdot} - \hat{\beta}_{0\cdot})^2$	—
within treatments	$np^s(r-1)$	$np^s(r-1)s_w^2$	σ^2
Total	np^sr-1	—	—

Thus, the test statistic (3.17) could test whether $\beta_l = 0$ or $\beta_l \neq 0$. It should be remarked that the F-like ratios (3.17) and (3.18) are not distributed like central or noncentral $F(v_1, v_2)$ random variables, because the distributions of these ratios, in the present case, are also affected by the variability introduced by the sampling procedure. The study of these distributions is reserved for another paper. The analysis of variance scheme suggested is summarized in Table 1. The sum of squares for between treatments serves, as usual, the purpose of simplifying the computation of the within sum of squares.

4. Randomization Procedure II in Fractional Replication Designs.

4.1 *The case of a n/p^{m-s} replicate of a p^m factorial experiment.* Randomization Procedures I and II differ substantially due to the fact that, in Randomization Procedure I, the confounded parameters are not the p^s chosen ones, while in Randomization Procedure II the confounded parameters are the p^s chosen ones. Thus, in Randomization Procedure II all the p^m treatment combinations are classified, into p^s blocks of equal size and from every block a random sample of n treatment combinations is drawn at random. While, in Randomization Procedure I the sampling of treatment combinations is essentially a cluster type of sampling, Randomization Procedure II is essentially a type of stratified sampling.

For the present Randomization Procedure we do not give a special presentation of the theory for the $n/2^{m-s}$ case, because the theory for this case can be derived in a manner similar to that of Section 3.1. Moreover, as will be seen later, the important results for a $n/2^{m-s}$ fractional replication can be derived directly from the results for the general case.

For the same reasons which were mentioned in Section 3.2, we require that the subgroup of chosen parameters be generated by s independent parameters $(\beta_{p^{i_0}}, \beta_{p^{i_1}}, \cdots, \beta_{p^{i_{s-1}}})$ representing main effects of s chosen factors.

There is no loss of generality if we assume that the chosen parameters are the first p^s ones, i.e., $(\beta_0, \beta_1, \cdots, \beta_{p^s-1})$. For this subgroup, of chosen parameters, the corresponding p^s blocks of treatment combinations are given by the sets

$$X_i = \{x : c_l^{(p^m)}(x_{i+jp^s}) = c_l^{(p^s)}(x_i) \tag{4.1}$$

for all $i = 0, \cdots, p^s - 1; j = 0, \cdots, p^{m-s} - 1$ and $l = 0, 1, \cdots, p^s - 1\}$.

Formula (4.1) results from the fact that here every block

$$X_i = (x_{i+jp^s} : i = 0, \cdots, p^s - 1; j = 0, \cdots, p^{m-s} - 1)$$

and from the equation $\mathbf{C}^{(p^m)} = \mathbf{C}^{(p^{m-s})} \otimes \mathbf{C}^{(p^s)}$. It should be remarked here again that the classification according to (4.1) is equivalent to the regular procedure of confounding the p^s chosen parameters.

From every block of treatment combinations X_i a random sample of $n (1 \leqq n \leqq p^{m-s} - 1)$ treatments is drawn. Let

$$S_i(x) = (x_{i+j_{i1}p^s}, x_{i+j_{i2}p^s}, \cdots, x_{i+j_{in}p^s})$$

be a random sample of treatment combinations from block X_i, and

$$\{y_{ij_k} : k = 1, \cdots, n\}$$

their associated random variables (treatment yields).

DEFINITION. An estimator of chosen p^s parameters is thus defined to be

$$\hat{\beta} = (\Delta^{(p^s)})^{-1}(\mathbf{C}^{(p^s)})' y, \tag{4.2}$$

where $y' = (y_{0\cdot}, y_{1\cdot}, \cdots, y_{(p^s-1)\cdot})$.

THEOREM 4.1: *For any given set of p^s samples of treatment combinations $S = \{S_i : i = 0, \cdots, p^s - 1\}$ the conditional expectation of an estimator*

$$\hat{\beta}_l \qquad (l = 0, \cdots, p^s - 1)$$

is given by

$$E(\hat{\beta}_l \mid S) = \beta_l + \frac{1}{d_l^{(p^s)}} \sum_{i=0}^{p^s-1} \sum_{r=0}^{p^s-1} c_{il}^{(p^s)} c_{ir}^{(p^s)} \left[\sum_{q=1}^{p^{m-s}-1} c_{j_i \cdot q}^{(p^{m-s})} \beta_{r+qp^s} \right], \tag{4.3}$$

where

$$c_{j_i \cdot q}^{(p^{(m-s)})} = n^{-1} \sum_{k=1}^{n} c_{j_{ikq}}^{(p^{m-s})}.$$

PROOF. According to the linear model (2.1),

$$E(y_{i\cdot} \mid S_i) = n^{-1} \sum_{k=1}^{n} \sum_{t=0}^{p^m-1} c_t^{(p^m)}(x_{i+j_{ik}p^s})\beta_t ,$$

where

$$c_t^{(p^m)}(x_{i+j_{ik}p^s}) = c_{i+j_{ik}p^s,t}^{(p^m)} ;$$

however, $c_t^{(p^m)}(x_{i+j_{ik}p^s}) = c_{it}^{(p^s)}$ for all $t = 0, 1, \cdots, p^s - 1$. Thus,

$$E(y_{i\cdot} \mid S_i) = \sum_{t=0}^{p^s-1} \beta_t c_{it}^{(p^s)} + n^{-1} \sum_{k=1}^{n} \sum_{t=p^s}^{p^m-1} \beta_t c_t^{(p^m)}(x_{i+j_{ik}p^s}).$$

Substituting

$$c_t^{(p^m)}(x_{i+j_{ik}p^s}) = c_{ir_t}^{(p^s)} \cdot c_{j_{ikq_t}}^{(p^{m-s})},$$

where

$$t = r_t + q_t p^s \qquad (q_t = 1, \cdots, p^{m-s} - 1; r_t = 0, \cdots, p^s - 1),$$

we obtain

$$E(y_{i\cdot} \mid S_i) = \sum_{t=0}^{p^s-1} c_{it}^{(p^s)} \beta_t + \sum_{q=1}^{p^{m-s}-1} c_{j_i \cdot q}^{(p^{m-s})} \sum_{r=0}^{p^s-1} c_{ir}^{(p^s)} \beta_{r+qp^s} .$$

By inserting this result into definition (4.2), we obtain formula (4.3).

THEOREM 4.2: *If the sampling of treatment combinations from every block X_i is at random, then β_l is an unbiased estimator of $\beta_l (l = 0, \cdots, p^s - 1)$ with variance*

$$V(\hat{\beta}_l) = \sigma^2 / n d_l^{(p^s)} + [M/np^{m-s}(d_l^{(p^s)})^2] \sum_{q=1}^{p^{m-s}-1} d_q^{(p^{m-s})} \sum_{i=0}^{p^s-1} (c_{il}^{(p^s)})^2 [c_{ir}^{(p^s)} \beta_{r+qp^s}]^2, \tag{4.4}$$

where

$$M = \begin{cases} 1, & \text{, if sampling is with replacement} \\ 1 - (n-1)/(p^{m-s}-1), & \text{if sampling is without replacement.} \end{cases}$$

PROOF.

(i) $\hat{\beta}_l$ is unbiased, because

$$E(\hat{\beta}_l) = E_S\{E(\hat{\beta}_l \mid S)\},$$

where $E_S(\,)$ is the expected value of the estimator in brackets, with respect to all the possible choices of samples $S = \{S_i : i = 0, \cdots, p^s - 1\}$. According to Theorem 4.1,

$$E(\hat{\beta}_l) = \beta_l + \left(\frac{1}{d_l^{(p^s)}}\right) \sum_{i=0}^{p^s-1} \sum_{r=0}^{p^s-1} c_{il}^{(p^s)} c_{ir}^{(p^s)} \sum_{q=1}^{p^{m-s}-1} \beta_{r+qp^s} E_{S_i}(c_{j_i \cdot q}^{(p^{m-s})}).$$

However, sampling is at random. Hence,

$$E_{S_i}(c_{j_i q}^{(p^{m-s})}) = \sum_{j=0}^{p^{m-s}-1} \frac{c_{jq}^{(p^{m-s})}}{p^{m-s}} = 0$$

for all $i = 0, \cdots, p^s - 1$ and all $q = 1, \cdots, p^{m-s} - 1$.

(ii) According to definition (4.2) of $\hat{\beta}_l$ and the linear model (2.1),

$$\hat{\beta}_l = \left(\frac{1}{d_l^{(p^s)}}\right) \sum_{i=0}^{p^s-1} c_{il}^{(p^s)} \left[\sum_{t=0}^{p^s-1} c_{it}^{(p^s)} \beta_t + n^{-1} \sum_{k=1}^{n} \sum_{t=p^s}^{p^m-1} c_{i+j_k p^s, t}^{(p^m)} \beta_t + \bar{e}_i \right],$$

where $V(e) = \sigma^2$ is independent of the x's. Moreover, samples from different blocks X_i are independent, thus,

$$V(\hat{\beta}_l) = \sigma^2 / n d_l^{(p^s)} + (d_l^{(p^s)})^{-2} \sum_{i=0}^{p^s-1} (c_{il}^{(p^s)})^2 V_{S_i} \left\{ n^{-1} \sum_{k=1}^{n} \sum_{t=p^s}^{p^m-1} c_{i+j_k p^s, t}^{(p^m)} \beta_t \right\},$$

where $V_{S_i}(\quad)$ is the variance of the estimator given inside the brackets $(\quad)$ over all the possible random samples S_i from block X_i.

$$V_{S_i}\left(n^{-1} \sum_{k=1}^{n} \sum_{t=p^s}^{p^m-1} c_{i+j_k p^s, t}^{(p^m)} \beta_t\right) = \sum_t \beta_t^2 V_{S_i}(n^{-1} \sum_k c_{i+j_k p^s, t}^{(p^m)})$$

$$+ \sum_{(t_1 \neq t_2) = p^s}^{p^m-1} \beta_{t_1} \beta_{t_2} \operatorname{cov}_{S_i}(n^{-1} \sum_k c_{i+j_k p^s, t_1}^{(p^m)}; n^{-1} \sum_k c_{i+j_k p^s, t_2}^{(p^m)}).$$

According to the lemma in Section 2, we can substitute

$$c_{i+j_k p^s, t}^{(p^m)} = c_{ir_t}^{(p^s)} c_{jkq_t}^{(p^{m-s})}, \qquad \text{where} \quad t = q_t p^s + r_t$$

$$q_t = 1, \cdots, p^{m-s} - 1 \quad \text{and} \quad r_t = 0, \cdots, p^s - 1.$$

Thus,

$$V_{S_i}\left(n^{-1} \sum_k \sum_t c_{i+j_k p^s, t}^{(p^m)} \beta_t\right) = \sum_{q=1}^{p^{m-s}-1} \sum_{r=0}^{p^s-1} \beta_{qp^s+r}^2 (c_{ir}^{(p^s)})^2 V_{S_i}\left(n^{-1} \sum_k c_{jkq}^{(p^{m-s})}\right)$$

$$+ \sum_{(q_1 \neq q_2)=1}^{p^{m-s}-1} \sum_{(r_1 \neq r_2)=0}^{p^s-1} \beta_{q_1 p^s + r_1} \beta_{q_2 p^s + r_2} c_{ir_1}^{(p^s)} c_{ir_2}^{(p^s)} \operatorname{cov}_{S_i}\left(n^{-1} \sum_k c_{jkq_1}^{(p^{m-s})}; n^{-1} \sum_k c_{jkq_2}^{(p^{m-s})}\right).$$

However,

$$\sum_{j=0}^{p^{m-s}-1} c_{jq}^{(p^{m-s})} = 0 \text{ for all } q = 1, 2, \cdots, p^{m-s} - 1.$$

Hence, according to the theory of random sampling,

$$V_{S_i}\left(n^{-1}\sum_k c_{jkq}^{(p^{m-s})}\right) = M\, d_q^{(p^{m-s})}/n\, p^{m-s},$$

where M is a "finite multiplier", defined in (4.4). Similarly,

$$\text{cov}_{S_i}\left(n^{-1}\sum_k c_{jkq_1}^{(p^{m-s})}; n^{-1}\sum_k c_{jkq_2}^{(p^{m-s})}\right) = \begin{cases} 0, & q_1 \neq q_2 \\ M\, d_q^{(p^{m-s})}/np^{m-s}, & q_1 = q_2. \end{cases}$$

It follows that

$$V_{S_i}\left(n^{-1}\sum_k\sum_t c_{i+jkp^s,t}^{(p^m)}\beta_t\right) = (M/np^{m-s})\sum_{q=1}^{p^{m-s}-1} d_q^{(p^{m-s})}\left[\sum_{r=0}^{p^s-1} c_{ir}^{(p^s)}\beta_{r+qp^s}\right]^2.$$

Substituting this result in the formula of $V(\hat{\beta}_l)$ yields formula (4.4). From this theorem the following corollaries are obtained:

(1) *In a $n/2^{m-s}$ fractional replication, according to Randomization Procedure II, all the variances of $\hat{\beta}_l (l = 0, \cdots, p^s - 1)$ are equal to*

$$V(\hat{\beta}_l) = \sigma^2/n2^s + (M/n2^s)\sum_{t=2^s}^{2^m-1}\beta_t^2. \tag{4.5}$$

(2) *In a $n/2^{m-s}$ fractional replication the variance of every estimator $\hat{\beta}_l (l = 0, \cdots, p^s - 1)$, according to Randomization Procedure II, is equal to the arithmetic mean of the variances of 2^s different estimators given by Randomization Procedure I.*

If we designate the variance of $\hat{\beta}_l$, according to Randomization Procedure I, by $V_{\text{I}}(\hat{\beta}_l)$ and that of $\hat{\beta}_l$, according to Randomization Procedure II, by $V_{\text{II}}(\hat{\beta}_l)$ then,

$$V_{\text{II}}(\hat{\beta}_l) = 2^{-s}\sum_{l=0}^{2^s-1} V_{\text{I}}(\hat{\beta}_l). \tag{4.6}$$

4.2. *Testing of hypotheses according to Randomization Procedure II.* In this section test statistics, appropriate for Randomization Procedure II, are suggested. The null hypotheses and alternatives are similar to those represented in Section 3.4.

There are some substantial differences between the analysis of variance for Randomization Procedure I and that of Randomization Procedure II. In the former, the significance of the nuisance parameters is tested by p^s different test statistics. In the latter, one tests the significance of all the nuisance parameters together. It will be shown also that in the case of $n/2^{m-s}$ fractional replication, the analysis of variance according to Randomization Procedure II might be more powerful than that of Randomization Procedure I. In order to make the analysis possible we require that the number of treatment combinations, chosen at random, with replacement, from every block X_i, $n \geqq 2$, and the number of repetitions of the chosen treatments $r \geqq 2$.

The total sum of squares about the grand mean is partitioned into three components. The first one measures the variability "within treatments", the second one measures the variability between choices within blocks, and the last one measures the variability between blocks. The analysis is similar to a nested classification type of design.

The expected value of the mean square between choices within blocks, MSC, shown in Table 2 is given by,

$$E(\text{MSC}) = \sigma^2 + r/p^m \sum_{q=1}^{p^{m-s}-1} d_q^{(p^{m-s})} \sum_{i=0}^{p^s-1} \left[\sum_{r=0}^{p^s-1} c_{ir}^{(p^s)} \beta_{r+qp^s} \right]^2. \tag{4.7}$$

Thus, in the case of a $n/2^{m-s}$ fractional replication,

$$E(\text{MSC}) = \sigma^2 + r \sum_{t=2^s}^{2^m-1} \beta_t^2 . \tag{4.8}$$

Comparing (4.8) to (4.5), we come to the conclusion that a proper test of the null hypothesis $H_0: \beta_l = 0 (l = 1. \cdots, 2^s - 1)$ against $H_1: \beta_l \neq 0$ when $p = 2$ is to compare the quadratic forms

$$Q^*(\hat{\beta}_{l.}) = nr2^s \hat{\beta}_{l.}^2. \qquad (l = 1, 2, \cdots, 2^s - 1) \tag{4.9}$$

with MSC. Thus, in the case of an $n/2^{m-s}$ fractional replication the appropriate analysis of variance that applies is given in Table 2.

TABLE 2

Analysis of Variance for Randomization Procedure II in the Case of an $r/2^{m-s}$ Fractional Replication

Source of Variation	D.F.	S.S.	M.S.	E(M.S.)
β_1	1	$nr2^s\hat{\beta}_{1.}^2$	$nr2^s\hat{\beta}_{1.}^2$	$\sigma^2 + r\sum_{t=2^s}^{2^m-1} \beta_t^2 + rn2^s\beta_1^2$
⋮	⋮	⋮	⋮	⋮
β_{2^s-1}	1	$nr2^s\hat{\beta}_{2^s-1.}^2$	$nr2^s\hat{\beta}_{2^s-1.}^2$	$\sigma^2 + r\sum_{t=2^s}^{2^m-1} \beta_t^2 + rn2^s\beta_{2^s-1}^2$
Between Blocks	$2^s - 1$	$nr\sum_{i=0}^{2^s-1} (y_{i\cdot} - \hat{\beta}_0)^2$	—	—
Between choices within blocks	$2^s(n-1)$	$r\sum_{i=0}^{2^s-1}\sum_{k=1}^{n} (y_{ijk} - y_{i\cdot})^2$	MSC	$\sigma^2 + r\sum_{t=2^s}^{2^m-1} \beta_t^2$
Within treatments	$2^sn(r-1)$	$\sum_{i=0}^{2^s-1}\sum_{k=1}^{n}\sum_{h=1}^{r} (y_{ijkh} - y_{ijk})^2$	s_w^2	σ^2
Total	$2^snr - 1$	—	—	—

Here, we test the significance of β_l by the F-like ratio

$$F_l^* = nr2^s\hat{\beta}_l^2/\text{MSC} \qquad (l = 1, 2, \cdots, 2^s - 1). \tag{4.10}$$

The significance of the nuisance parameters is tested by the F-like ratio $F_c^* = \text{MSC}/s_w^2$. The distribution function of F_c^* is the average, over all the $\left[\binom{2^{m-s}}{n}\right]^{2^s}$ possible samples, of non-central-F distribution functions. The distribution function of F_l^* is more complicated.

In the case of a n/p^{m-s} $(p \geqq 3)$ fractional replication we suggest testing the hypothesis $H_0:\beta_l = 0$ against $H_1:\beta_l \neq 0$ for every $l = 1, \cdots, p^s - 1$ independently by an analysis of variance similar to that given for Randomization Procedure I. We first sample one treatment combination, at random, from every block X_i independently and from the obtained set of random variables $(y_{0j_1}, y_{1j_1}, \cdots, y_{p^s-1,j_1})$ estimate the chosen parameters. Call this vector $\hat{\beta}_1$. In a similar manner we repeat this sampling procedure and estimation n times $(n \geqq 2)$. Thus, for every parameter $\hat{\beta}_l$ $(l = 0, \cdots, p^s - 1)$ we obtain a set of n estimates $(\hat{\beta}_{l_1}, \cdots, \hat{\beta}_{l_n})$.

Define, for every $l = 0, \cdots, p^s - 1$,

$$s_{\hat{\beta}_l}^2 = (n-1)^{-1}\sum_{k=1}^{n}(\hat{\beta}_{l_k} - \hat{\beta}_{l\cdot})^2, \qquad \text{where } \hat{\beta}_{l\cdot} = n^{-1}\sum_{k=1}^{n}\hat{\beta}_{l_k}. \tag{4.11}$$

Because sampling is at random with replacement, it follows that $E(s_{\hat{\beta}_l}^2) = V(\hat{\beta}_l)$. Hence, a proper test statistic for $H_0:\beta_l = 0.$ against $H_1:\beta_l \neq 0$ is

$$F_l^* = np^s\hat{\beta}_{l\cdot}^2/s_{\hat{\beta}_l}^2, \tag{4.12}$$

with $f_1 = 1$ and $f_2 = n - 1$ degrees of freedom.

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
Chosen Parameters	(M)	(A)	(A^2)	(B)	(AB)	(A^2B)	(B^2)	(AB^2)	(A^2B^2)
true values	198.5	39.8	−68.5	33.9	46.4	21.0	18.4	−19.0	3.1
	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	β_{16}	β_{17}
aliases with respect to C	(C)	(AC)	(A^2C)	(BC)	(ABC)	(A^2BC)	(B^2C)	(AB^2C)	(A^2B^2C)
true value	13.8	−21.2	−22.3	2.5	−15.6	−12.1	−9.7	9.0	5.5
	β_{18}	β_{19}	β_{20}	β_{21}	β_{22}	β_{23}	β_{24}	β_{25}	β_{26}
aliases with respect to C^2	(C^2)	(AC^2)	(A^2C^2)	(BC^2)	(ABC^2)	(A^2BC^2)	B^2C^2	(AB^2C^2)	$(A^2B^2C^2)$
True value	−10.9	7.4	−2.7	5.6	−11.0	−6.7	3.4	5.7	−2.5

In (4.7) we see a possible test statistic for testing the hypothesis

$$H_0 : \text{All } \beta_t = 0 \text{ for } t = p^s, p^s + 1, \cdots, p^m - 1$$

against

$$H_1 : \text{At least one parameter } \beta_t \ (t = p^s, \cdots, p^m - 1) \text{ is not zero,}$$

is $F_c^* = \text{MSC}/s_w^2$ with $f_1 = p^s(n-1)$ and $f_2 = p^s n(r-1)$ degrees of freedom.

EXAMPLE. In the following we illustrate the possible effects of the two randomization procedures, by an example of a 3^3 factorial system given by O. L. Davies, ([6], p. 353). The three factors studied are designated by A, B and C. The chosen parameters are those generated by the main effects A and B. A $\frac{2}{3}$ fractional replication is considered. For the purposes of illustration we assume that the estimates given by Davies are true values. (See page 290.)

The standard deviation of Y, for any treatment combination is $\sigma^2 = 27$. In the case of fixed fractional replication designs, the biases and standard errors of the estimators of the chosen parameters are given in the following table: (Confounding here is according to M, C, and C^2.)

Chosen parameters	True value	Bias			Standard Error
		Fixed Design 1	Fixed Design 2	Fixed Design 3	
M	198.5	0.316	−10.344	10.028	6.364
A	39.8	4.250	6.567	−10.817	7.794
A^2	−68.5	13.383	−2.422	−10.961	4.500
B	33.9	−1.400	5.600	−4.200	7.794
AB	46.4	4.425	−11.000	6.575	9.546
A^2B	21.0	10.725	−6.700	−4.025	5.511
B^2	18.4	4.033	2.844	−6.877	4.500
AB^2	−19.0	−8.675	6.533	2.142	5.511
A^2B^2	3.1	−1.058	−2.778	3.836	3.182

Fixed Design 1, 2 and 3 are taken to be the blocks (X_0, X_1), (X_0, X_2) and (X_1, X_2) respectively. With the randomization procedure there is no bias but the variance of the estimators are increased. Standard errors of the estimators for both procedures are given below.

Parameters	Rand. Procedure I		Rand. Procedure II		Fixed (without Rand.)
	With replacement	Without replacement	With replacement	Without replacement	
M	14.9	11.5	12.9	10.2	6.4
A	16.3	12.8	13.3	10.9	7.8
A^2	13.9	10.3	10.3	8.0	4.5
B	9.7	8.8	13.9	11.3	7.8
AB	17.1	13.9	16.1	13.2	9.5
A^2B	11.1	8.8	8.9	7.4	5.5
B^2	7.9	6.5	8.8	7.0	4.5
AB^2	9.5	7.8	9.0	7.5	5.5
A^2B^2	5.0	4.3	8.6	6.5	3.2

It was shown in Sections 3 and 4 that, for the sake of testing hypotheses, sampling should be done with replacement. In this example, variances of estimators given by randomization procedures are about twice as large as those given by the fixed procedure (without randomization). However, the presence of bias of the fixed procedure gives sometimes benefit to estimators of the randomization procedure. This will be discussed fully in the next section.

5. Discussion. In this section we comment on various questions which arise in connection with the paper.

1. It is apparent from the development in Sections 3 and 4 that both randomization procedures can readily be applied using standard confounding methods, see [3], [5], [6], [9], [15]. The use of the matrices $\mathbf{C}^{(p^m)}$ is particularly convenient since they can readily be written down and confounding only involves looking at suitable columns of the matrices. Similarly, the application of the analysis of variance is straight-forward.

2. In order to compare Procedures I and II let us first consider the case of the 2^m factorial system. In Procedure II the variances of the estimate of the parameters of interest are constant, while in procedure I the variances may not be. The relationship of the variance for the two methods is given in the corollary at the end of Section 4.1. If no information is available concerning nuisance parameters procedure II seems preferable since it guards against excesses in variance. This is particularly true if one is equally interested in the parameters of interest. However, if information concerning "nuisance" parameters is available it might, with profit, be used to choose the defining parameters in Procedure I so that the variances of particular parameters of interest are reduced. This, of course, takes place at the expense of increasing other variances. This, can be particularly useful if one does not have equal interest in the parameters of interest. When $p \geqq 3$, the comparison of the two procedures becomes more complicated. However, here again, Procedure II takes into account all the nuisance parameters, while Procedure I only the aliases of each parameter.

Another aspect, in the comparison of Procedures I and II, for the case of the 2^m, is the respective abilities for testing the significance of parameters of interest with the two methods. In Procedure I, the test is made in terms of an F-like ratio with degrees of freedom $f_1 = 1$ and $f_2 = n - 1$ while in Procedure II we have $f_1 = 1$ and $f_2 = 2^s(n - 1)$.

A further comparison between Procedures I and II is in the respective methods for testing the significance of the "nuisance" parameters. In Procedure I one tests the nuisance parameters in blocks of aliases, while in Procedure II the test is in terms of *all* the nuisance parameters simultaneously. The degrees of freedom, for error, in Procedure II is, however, larger. It is clear that the relative merits of the two procedures depends on the purpose of the experiment. In the case where randomized factorial experiments are used, for exploratory purposes, it seems a definite advantage to be able to test the nuisance parameters in blocks. This is so, since such tests may shed light on how to proceed further.

3. The question of confidence intervals for the parameters of interest can be approached in at least two ways, according to whether information about nuisance parameters is or is not available.

One approach is to use the confidence interval suggested by the analysis of variance tables as if the usual quantities had t-distributions. There is reason to believe that this approach has merit, since the t-distribution is known to be robust against departures from normality. The adequacy of this approach will also depend on σ^2 and the nuisance parameters. These questions will be further treated, as part of the general distribution problems arising from the randomization procedures, in a subsequent paper.

Another approach depends on some knowledge of the nuisance parameters. Let β_l ($l = 0, \cdots, p^s - 1$) be an estimated parameter and $\hat{\beta}_l$ its unbiased estimator. What is the interval $(\beta_l - \epsilon, \beta_l + \epsilon)$ for which the probability is at least $1 - \alpha$ that the estimated value belongs to that interval?

The conditional distribution function of $\hat{\beta}_l$ given a block or a sample of treatment combinations, X, is normal with conditional mean $E(\hat{\beta}_l \mid X)$ and standard deviation $\sigma/(nr\, d_l^{(p^s)})^{\frac{1}{2}}$. Thus

$$
\begin{aligned}
P\{\beta_l - \epsilon \leqq \hat{\beta}_l \leqq \beta_l + \epsilon\} &= E_x\{P(\beta_l - \epsilon \leqq \hat{\beta}_l \leqq \beta_l + \epsilon \mid X)\} \\
&= E_X\left\{\Phi\left(\frac{\beta_l + \epsilon - E(\hat{\beta}_l \mid X)}{\sigma/(nr\, d_l^{(p^s)})^{\frac{1}{2}}}\right)\right\} - E_X\left\{\Phi\left(\frac{\beta_l - \epsilon - E(\hat{\beta}_l \mid X)}{\sigma/(nr\, d_l^{(p^s)})^{\frac{1}{2}}}\right)\right\},
\end{aligned} \tag{5.1}
$$

where $\Phi(u)$ is the cumulative normal distribution function with zero mean and unit variance.

Let $E^*(\hat{\beta}_l \mid X) = E(\hat{\beta}_l \mid X) - \beta_l$. Explicit formulae for $E(\hat{\beta}_l \mid X)$ are given by (3.17) and (4.3).

Let us expand the function $\Phi(u)$ into a power series about $u = \epsilon(nr\, d_l^{(p^s)})^{\frac{1}{2}} / \sigma$. We obtain

$$
P\{\beta_l - \epsilon \leqq \hat{\beta}_l \leqq \beta_l + \epsilon\} = \sum_{j=0}^{\infty} \frac{2E_X\{(E^*(\hat{\beta}_l \mid X))^{2j}\}}{\epsilon^{2j}(2j)!}\, u^{2j}\Phi^{(2j)}(u) - 1, \tag{5.2}
$$

where $\Phi^{(j)}(u)$ is the jth order derivative of $\Phi(u)$.

A fairly good approximation is obtained if only the first two terms ($j = 0, 1$) of the above series are considered. Thus,

$$
\begin{aligned}
&P\{\beta_l - \epsilon \leqq \hat{\beta}_l \leqq \beta_l + \epsilon\} \\
&\qquad \cong 2\Phi(u) - 1 + \frac{u^2}{\epsilon^2}\Phi^{(2)}(u)E_X\{(E^*(\hat{\beta}_l \mid X))^2\}.
\end{aligned} \tag{5.3}
$$

Let

$$
D_l = (V(\hat{\beta}_l) - \sigma^2/nr\, d_l^{(p^s)})(nr\, d_l^{(p^s)}/\sigma^2), \tag{5.4}
$$

where $V(\hat{\beta}_l)$ is given either by Randomization Procedure I or II. Formula (5.3), after some manipulation, reduces to

$$\text{(5.5)} \qquad \text{Prob}\{\beta_l - \epsilon \leqq \hat{\beta}_l \leqq \beta_l + \epsilon\} \cong 2\Phi(u) - 1 - D_{l\varphi}(u),$$

where $\varphi(u)$ is the normal density and $u = \epsilon(nr\, d_l^{(p^s)})^{\frac{1}{2}}/\sigma$.

The quantity D_l depends on the nuisance parameters and is a measure of the excess of variability due to randomization. If all the nuisance parameters or (for Procedure I) only the alias parameters are zero, then $D_l = 0$.

It is clear that in order to obtain a probability $1 - \alpha$ the value of ϵ is given, approximately, by the root of the equation

$$\text{(5.6)} \qquad \Phi(u) - \varphi(u)uD_l/2 = 1 - \alpha/2.$$

The relationship between the required value of ϵ, the confidence level $1 - \alpha$ and the measure of excess due to randomization D_l can easily be represented graphically. It is clear that the application of the method depends only on having an upper bound for D_l. For a numerical illustration, we return to the example previously presented. The values represented in the following table measure half of the length of 0.95 confidence intervals of β_l for sampling with replacement.

Chosen Parameters	True Value	Rand. Proced. I	Rand. Proced. II	When $D_l = 0$
M	198.5	19.84	18.88	13.10
A	39.8	23.38	21.82	16.10
A^2	−68.5	14.85	13.72	9.18
B	33.9	18.71	21.43	16.10
AB	46.4	26.73	26.25	19.50
A^2B	21.0	16.53	14.60	11.30
B^2	18.4	12.82	13.10	9.18
AB^2	−19.0	12.81	14.93	11.30
A^2B^2	3.1	8.27	10.18	6.50

With respect to the length of the 0.95 confidence interval, Randomization Procedure II is slightly better than Randomization Procedure I. In Randomization Procedure II the maximum length over all the intervals is 26.25 while in Randomization Procedure I the maximum length is 26.73. When all the nuisance parameters are zero ($D_l = 0$ for all $l = 0, \cdots, p^s - 1$), all the confidence intervals are uniformly shorter.

4. The study of the distribution functions of the test statistics (3.17), (3.18), (4.13) and (4.14) is very important for the determination of the level of significance and power function, in the analysis of variance. It can be readily shown that the distribution under the null hypothesis, of test statistics (3.18) and (4.14), which tests the significance of the nuisance parameters, is like that of a central F, with $f_1 = n - 1$ and $f_2 = np^s(r - 1)$ degrees of freedom. However, under the alternative hypothesis, when $D_l > 0$, the distribution functions of these test statistics is the average, over all the possible samples, of non-central F distribution functions with f_1 and f_2 degrees of freedom and of non centrality:

$$\text{(5.9)} \qquad \lambda_{l\cdot}(S) = (r\, d_l^{(p^s)}/2\sigma^2)\sum_{j=1}^{n}[E^*(\hat{\beta}_{lw_j} \mid S) - E^*(\hat{\beta}_{l\cdot} \mid S)]^2$$

where $E^*(\hat{\beta}_{l\cdot} \mid S) = n^{-1}\sum_{j=1}^{n} E^*(\hat{\beta}_{lv_j} \mid S)$. These distribution functions are given approximately by

$$G(F \mid f_1, f_2; \lambda_l) \cong H(F \mid f_1, f_2; \Lambda_{l\cdot}) + \tfrac{1}{2} V_S(\lambda_{l\cdot}(S)) \left[\sum_{j=0}^{2} (-1)^j \binom{2}{j} H\left(F \frac{f_1}{f_1 + 4 - 2j} \,\middle|\, f_1 + 2j, f_2; \Lambda_{l\cdot}\right)\right], \tag{5.10}$$

where $H(F \mid f_1, f_2; \Lambda_{l\cdot})$ is a non central F distribution with f_1 and f_2 degrees of freedom, and parameter of non-centrality

$$\Lambda_{l\cdot} = E_S(\lambda_{l\cdot}(S)) = f_1 D_l/2. \tag{5.11}$$

The distribution of test statistics (3.17) and (4.13), which test the significance of the interesting parameters, is more complicated. Its conditional distributions, for any given sample is like that of the ratio of two noncentral chi-squares. Under the null-hypothesis these distribution functions are given approximately by:

$$\begin{aligned} G(F^*) &\cong H(F^* \mid 1, n-1; D_l/2; D_l(n-1)/2) \\ &+ D_l^2/4\left[\sum_{j=0}^{2} (-1)^j \binom{2}{j} H\left(F^* \frac{1}{5-2j} \,\middle|\, 5-2j, n-1; \frac{D_l}{2}; \frac{(n-1)D_l}{2}\right)\right] \\ &+ (n-1)D_l^2/4\left[\sum_{j=0}^{2} (-1)^j \binom{2}{j}\right. \\ &\left. \cdot H\left(F^* \frac{(n-1)+4-2j}{n-1} \,\middle|\, 1, n+3-2j; \frac{D_l}{2}; \frac{(n-1)D_l}{2}\right)\right], \end{aligned} \tag{5.12}$$

where

$$H(x \mid f_1, f_2; \alpha, \beta) = \sum_{j=0}^{\infty} e^{-\beta} \frac{\beta^j}{j!} H\left(x \frac{f_2 + 2j}{f_2} \,\middle|\, f_1, f_2; \alpha\right). \tag{5.13}$$

The validity of these approximations, and the determination of test criterions for a given level of significance will be given elsewhere. Numerical computations indicate that $G(F^*)$, in many circumstances, can be approximated adequately by the central F distribution.

5. One of the relevant aspects in the comparison between the randomization procedures and the "classical" fractional replication designs is that the classical design may give biased estimates. The randomization designs give unbiased estimates with variances, say V_{R_i}, while the classical designs give possible biased estimates with variance V_c and bias, say B. In general, we have that $V_{R_i} \geqq V_c$. A relevant factor in the comparison between randomization and nonrandomization is the old problem of comparing variance and bias. In a sense the randomization removes bias at the expense of increased variance. How should one compare $(V_{R_i}, 0)$ and (V_c, B)? This is a variant of the problem of balancing accuracy and precision in measurement. On the one hand, it is clear that it is useless to have a very precise inaccurate estimate, on the other hand we do not want an accurate but very imprecise estimate. There are at least two approaches to this

problem of adopting an appropriate criterion. One criterion is simply to look at the variance plus the bias squared. In other words, choose the procedure, p, to minimize $V_p + B_p^2$, where V_p is the variance using procedure p and B_p is the bias using procedure p. This criterion was adopted by G.E.P. Box [2].

Another approach is to adopt a "closeness" criterion. Suppose we are comparing $(V_R, 0)$ with (V_c, B). Let us compute, Prob $\{|\hat{\beta} - \beta| \leqq \lambda|\beta|\}$ for a particular procedure, where $\hat{\beta}$ is the estimator using the procedure and $0 \leqq \lambda \leqq 1$. The parameter λ measures how important it is to be close to β.

We give, below, some calculations of the probability, for $\lambda = 0.2$, for the previously discussed example.

Parameters	closeness = Prob$\{\|\hat{\beta} - \beta\| \leqq \lambda\|\beta\|\}$				
	Fixed Design 1	Fixed Design 2	Fixed Design 3	Rand. Proced. I (without replacement)	Rand. Proced. II
M	0.858	0.999	0.999	0.952	0.999
A	0.229	0.540	0.348	0.372	0.458
A^2	0.348	0.994	0.729	0.690	0.965
B	0.197	0.504	0.550	0.417	0.399
AB	0.266	0.412	0.562	0.413	0.467
A^2B	0.114	0.303	0.444	0.287	0.368
B^2	0.108	0.502	0.230	0.280	0.263
AB^2	0.106	0.298	0.478	0.294	0.319
A^2B^2	0.008	0.106	0.075	0.063	0.001

This table illustrates that for 6 out of the 9 chosen parameters, Randomization Procedure II is better than Randomization Procedure I, with respect to the closeness criterion. Fixed Design 2 is almost always better than both Randomization Procedures. However, Fixed Design 1 is always the worst. If information about the nuisance parameters is not available, there is no way how to decide which Fixed Design is a good one and which is a bad one. Thus, Randomization Procedures I and II guard against a bad choice of a design when the nuisance parameters are unknown. Finally, it is to be emphasized that fixed fractional replication, or for that matter a full replicate of a factorial design, requires assumptions about parameters, usually of the form that high order interactions are negligible. However, for the randomization schemes, for $n \geqq 2$, no such assumptions are required.

REFERENCES

[1] R. C. Bose, "Mathematical theory of the symmetrical factorial design," *Sankhyā*, Vol. 8 (1947), pp. 107–166.

[2] G. E. P. Box and N. R. Draper, "A basis for the selection of a response surface design," *J. Amer. Stat. Assn.*, Vol. 54 (1959), pp. 622–654.

[3] W. G. Cochran and G. M. Cox, *Experimental Design*, 2nd. ed., John Wiley and Sons, New York, 1957.

[4] J. CORNFIELD AND J. W. TUKEY, "Average values of mean squares in factorials," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 907–909.

[5] C. DANIEL, "Fractional replication in industrial research," *Third Berkeley Symposium on Probability and Statistics*, Vol. 5, University of California Press, Berkeley, 1956.

[6] O. L. DAVIES, ed., *Design and Analysis of Industrial Experiments*, Oliver and Boyd, London and Edinburgh, 1954.

[7] R. A. FISHER, *The Design of Experiments*, 4th ed., Oliver and Boyd, Edinburgh, 1947.

[8] O. KEMPTHORNE, *The Design and Analysis of Experiments*, John Wiley and Sons, New York, 1952.

[9] O. KEMPTHORNE, "The randomization theory of experimental inference," *J. Amer. Stat. Assn.*, Vol. 50 (1955), pp. 946–967.

[10] E. J. G. PITMAN, "Significance tests which may be applied to samples from any population, III, The analysis of variance test," *Biometrika*, Vol. 29 (1937), pp. 322–335.

[11] F. E. SATTERTHWAITE, "Random balance experiments," *Technometrics*, Vol. 1 (1959), pp. 111–138.

[12] B. L. WELCH, "On the Z-test in randomized blocks and Latin squares," *Biometrika*, Vol. 29 (1937), pp. 21–52.

[13] M. B. WILK AND O. KEMPTHORNE, "Some aspects of the analysis of factorial experiments in a completely randomized design," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 950–985.

[14] A. WALD AND J. WOLFOWITZ, "Statistical tests based on permutations of the observations," *Ann. Math Stat.*, Vol. 25 (1954), pp. 358–372.

[15] F. YATES, *The Design and Analysis of Factorial Experiments*, Imperial Bureau of Soil Science, Harpenden, England, 1937.