UNCLASSIFIED

AD **402 644**

Reproduced
by the

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA

UNCLASSIFIED

402 644

$6\underline{3}\underline{3}\ 3$

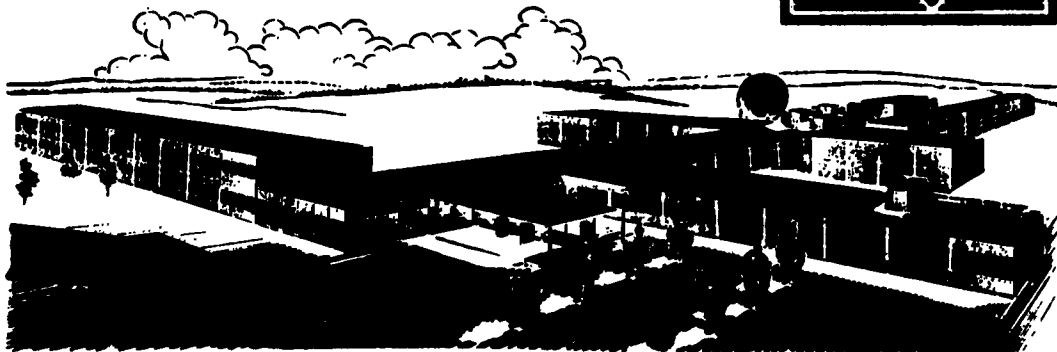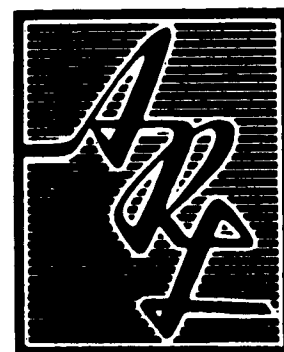# SOME ASPECTS OF CONSTRAINED RANDOMIZATION

GARY J. SUTTER
GEORGE ZYSKIND
OSCAR KEMPTHORNE

IOWA STATE UNIVERSITY
AMES, IOWA

ASTIA

APR 30 1963

JANUARY 1963

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

- - - - - - - - - -

Qualified requesters may obtain copies of this report from the Armed Services Technical Information Agency, (ASTIA), Arlington Hall Station, Arlington 12, Virginia.

- - - - - - - - - -

This report has been released to the Office of Technical Services, U. S. Department of Commerce, Washington 25, D. C. for sale to the general public.

- - - - - - - - - -

Copies of ARL Technical Documentary Reports should not be returned to Aeronautical Research Laboratory unless return is required by security considerations, contractual obligations, or notices on a specific document.

Aeronautical Research Laboratories,Wright-
Patterson AFB, Ohio. SOME ASPECTS OF CON-
STRAINED RANDOMIZATION, by Gary J. Sutter,
George Zyskind, Oscar Kempthorne, Iowa State
University, Ames, Iowa. November 1962. 30p.
incl. tables (Project 7071; Task 7071-01).
(Contract AF33(616)-8269) (ARL

Unclassified Report

In Chapter I the definition of con-
strained sets of plans and the problems to
be considered are given. In Chapter II a
correspondence between constrained sets of
plans and RT x RT square arrays and a corre-
spondence between constrained sets of plans
and resolvable balanced incomplete block

(over)

designs are given. These correspondences
led to the enumeration of many constrained
designs. In Chapter II some relationships
between the variance of the error sum of
squares under complete randomization and
under constrained randomization are given.
In Chapter IV the power of the randomization
test under constrained randomization was
considered for nearly linearly, randomly,
and semi-randomly distributed basal yields.
The empirical results exhibit a tendency of
the sensitivity of the test for the detection
of treatment differences to increase with
decreases in the variance of the error sum
of squares of the constrained sets.

SOME ASPECTS OF CONSTRAINED RANDOMIZATION

Gary J. Sutter

George Zyskind

Oscar Kempthorne

Iowa State University

Ames, Iowa

JANUARY 1963

FOREWORD

This interim technical report was prepared by Iowa State University, Ames, Iowa, for the Aeronautical Research Laboratories, Office of Aerospace Research, United States Air Force. The work reported herein was accomplished on Task 7071-01, "Research in Mathematical Statistics and Probability Theory" of Project 7071, "Mathematical Techniques of Aeromechanics".

ABSTRACT

In Chapter I the definition of constrained sets of plans and the problems to be considered are given. In Chapter II a correspondence between constrained sets of plans and $RT \times RT$ square arrays and a correspondence between constrained sets of plans and resolvable balanced incomplete block designs are given. These correspondences led to the enumeration of many constrained designs. In Chapter III some relationships between the variance of the error sum of squares under complete randomization and under constrained randomization are given. In Chapter IV the power of the randomization test under constrained randomization was considered for nearly linearly, randomly, and semi-randomly distributed basal yields. The empirical results exhibit a tendency of the sensitivity of the test for the detection of treatment differences to increase with decreases in the variance of the error sum of squares of the constrained sets.

TABLE OF CONTENTS

I. INTRODUCTION

The notion of constrained randomization is due to W. J. Youden (1956). Suppose that one has RT experimental units on which to apply T treatments. One might, for example, be comparing 3 measuring devices (T = 3 "treatments") by making 2 measurements with each, giving 6 measurements in all. The measurements will be taken in 6 consecutive periods, and one might be concerned about the possibility of drift in the attribute being measured or in the operators. One would then regard the 6 periods as experimental units, and choose a pattern of observation by associating devices to periods at random, subject to the restriction that each device is used in two periods. Such a way of proceeding is known as a completely randomized design. With this design some of the $\frac{(RT)!}{(R!)^T} \left[= 90 \right]$ possible plans have a systematic ordering of the treatments. Note that these occur in sets of T!, which differ only in the labelling of the treatments. With the case T = 3, R = 2, a possible plan under complete randomization is

Unit	1	2	3	4	5	6
Treatment	a	a	b	b	c	c

where the numbers 1, 2, ..., 6 denote the units and the letters a, b, c denote the treatments. Youden's contribution was the realization that there exists a subset of the totality of possible plans which has those properties of the complete set which are commonly regarded as essential, and does not contain some plans which seem to be undesirable because they appear to be systematic.

1

Definition: A subset of plans is an unbiased subset of plans if
(a) the probability that any particular treatment falls
on a particular unit is $\frac{1}{T}$, and the treatments occur
equally frequently in each plan, and (b) the expec-
tations of the treatment mean square and of the error
mean square over this subset are equal in the absence
of treatment effects.

The above implies that if the experiment were a dummy one, that is,
there were no treatment effects and no error of observation, the mean
square within treatments should equal the mean square between treat-
ments on the average. In this case either would then necessarily be
equal to

$$\frac{1}{RT-1} \sum_{i=1}^{RT} (y_i - \bar{y})^2 = \frac{1}{2RT(RT-1)} \sum_{\substack{i, j=1 \\ i \neq j}}^{RT} (y_i - y_j)^2$$

where the y_i $(i = 1, 2, \ldots, RT)$ are the basal yields of the experimental
units. Therefore the subset of plans is unbiased if all possible
differences between pairs of experimental units are represented in the
sum of squares for error equally frequently over all the plans belonging
to the subset.

Definition: A constrained set of plans is a subset of the totality
of plans which is unbiased.

The problems to be considered in this report are the existence and
properties of constrained sets.

2

II. EXISTENCE OF CONSTRAINED SETS OF PLANS

A. Representation in RT x RT Square Arrays

Consider the RT x RT square array of $(RT)^2$ cells (i, j), $(i, j = 1, 2, \ldots, RT)$ where the cell (i, j) is the intersection of the i-th row and j-th column in the array. Every time a particular treatment occurs on the i-th and j-th units of some plan, say plan I, we will place the symbol I in the (i, j)-th cell of the RT x RT square array. The symbol I in the (i, j)-th cell also corresponds to the $(y_i - y_j)^2$ component of the error mean square appearing in plan I. Since there are no differences of the form $(y_i - y_i)$, the diagonal cells of the square array never contain any symbols. Also since

$$\frac{1}{2RT(RT-1)} \sum_{\substack{i, j=1 \\ i \neq j}}^{RT} (y_i - y_j)^2 = \frac{1}{RT(RT-1)} \sum_{\substack{i, j=1 \\ i < j}}^{RT} (y_i - y_j)^2 ,$$

we need only consider the cells (i, j) for $i < j$ (i. e. the cells above the main diagonal). Consider, for example, the following constrained set which is not necessarily a useful one, for 3 treatments a, b and c, on 9 units:

Plan No.	Unit No.								
	1	2	3	4	5	6	7	8	9
A	a	a	a	b	b	b	c	c	c
B	a	b	c	a	b	c	a	b	c
C	a	b	c	c	a	b	b	c	a
D	a	b	c	b	c	a	c	a	b

3

The occurrence of differences between units in the error sums of squares can be represented by the following array

	1	2	3	4	5	6	7	8	9
1	O	A	A	B	C	D	B	D	C
2	A	O	A	D	B	C	C	B	D
3	A	A	O	C	D	B	D	C	B
4	B	D	C	O	A	A	B	C	D
5	C	B	D	A	O	A	D	B	C
6	D	C	B	A	A	O	C	D	B
7	B	C	D	B	D	C	O	A	A
8	D	B	C	C	B	D	A	O	A
9	C	D	B	D	C	B	A	A	O

For instance, the differences of units 1 and 2, units 1 and 3, units 2 and 3, etc. occur in the error sum of squares for plan A. Consideration of this representation leads to the following obvious theorem:

Theorem 1:

A necessary and sufficient condition for the existence of a constrained set of, say, N plans is that we can insert symmetrically the N symbols, $A, B, C, \ldots, N$ in the off-diagonal cells of the $RT \times RT$ array in such a way that

(a) any symbol occurs once or not at all in any cell,

(b) if a symbol occurs in cells (i, j) and (i, j') it occurs in cell (j, j'),

and (c) each cell contains the same number of symbols.

4

/

The number of cells in the upper off-diagonal portion of the RT x RT

square array is $\frac{RT(RT-1)}{2}$. Each plan has $\begin{pmatrix} R \\ 2 \end{pmatrix}$ $T = \frac{RT(R-1)}{2}$

differences between units treated alike so that the minimum possible

number of plans belonging to a constrained set is

$$N = \frac{RT(RT-1)}{2} \cdot \frac{2}{RT(R-1)} K = \frac{(RT-1)}{(R-1)} K , \qquad (1)$$

where K is the lowest integer such that N is an integer. It is obvious

then that K is the number of symbols contained in the (i, j)-th cell for

$i, j = 1, 2, \ldots, RT,$ $i \neq j,$ or the number of times the experimental unit

difference $(y_i - y_j)$ for $i, j = 1, 2, \ldots, RT,$ $i \neq j,$ occurs in the set of

plans.

B. Correspondence to Resolvable Balanced Incomplete Block Designs

We now make a correspondence between constrained sets of plans

and resolvable balanced incomplete block designs which we will hence-

forth designate by R. B. I. B. D.

A balanced incomplete block design with parameters $v, b, r, k,$ and

λ is an arrangement of v varieties in b blocks of k units with every

variety occurring in r blocks and every two varieties occurring

together in λ blocks. If in addition it is possible to arrange the b

blocks of the balanced incomplete block design into r groups of n

blocks each $(b = nr)$, so that each variety occurs exactly once in each

group of blocks, i. e. each group contains one complete replication of all

the varieties, then the design is a resolvable balanced incomplete block

design (R.B.I.B.D.).

Let S_i $(i = 1, 2, \ldots, N)$ denote the i-th group of blocks and

B_{ij} $(i = 1, 2, \ldots, N;$ $j = 1, 2, \ldots, T)$ denote the j-th block belonging to

S_i. Let the v varieties belonging to S_i correspond to the RT experi-

mental units of the i-th plan belonging to the constrained set. Let the

$k(= R)$ varieties of B_{ij} correspond to the R experimental units on

which treatment j is replicated in the i-th plan belonging to the

constrained set. By this correspondence we have determined that the

parameters of the R.B.I.B.D. are given by

$$v = RT$$
$$b = NT$$
$$r = N \tag{2}$$
$$k = R$$
$$\lambda = \frac{r(k-1)}{(v-1)} = \frac{N(R-1)}{(RT-1)} .$$

As an example of this correspondence consider the R.B.I.B.D. with

the parameters $v = 6$, $b = 15$, $r = 5$, $k = 2$, and $\lambda = 1$ which corresponds

to the constrained set of 5 plans for 3 treatments on 6 experimental

units. A R.B.I.B.D. with these parameters is given by

12	13	14	15	16
35	26	23	24	25
46	45	56	36	34

where the vertical lines separate the groups of blocks. From this

R.B.I.B.D. we obtain the constrained set of plans

Plan No.	Unit No.					
	1	2	3	4	5	6
I	a	a	b	c	b	c
II	a	b	a	c	c	b
III	a	b	b	a	c	c
IV	a	b	c	b	a	c
V	a	b	c	c	b	a

Thus for the first replicate of the R.B.I.B.D., the first block contains treatments 1 and 2, the second block contains treatments 3 and 5, and the third block contains treatments 4 and 6. The way of making the correspondence to constrained designs leads to the first plan having the first treatment (labelled a) being on units 1 and 2, the second treatment (labelled b) being on units 3 and 5 and the third treatment (labelled c) being on units 4 and 6. By constructing the RT x RT square array this set of plans may easily be shown to form a constrained set of plans.

Theorem 2:

A necessary and sufficient condition for the existence of a constrained set of N plans for T treatments on RT experimental units is the existence of a R.B.I.B.D. with parameters given by equation 2.

One form of the problem of finding a constrained set reduces therefore to the problem of finding R.B.I.B.D. with parameters given by equation 2. A survey of literature is given in Sutter (1962) and the specific references concerning the construction of R.B.I.B.D.'s are given at the end of this report.

III. PROPERTIES OF THE ERROR SUM OF SQUARES

A. The Error Sum of Squares under Complete Randomization

In the case of the completely randomized design for T treatments applied to RT experimental units let δ_u^{kf} be unity if the f-th realization of treatment k appears on the u-th unit and zero otherwise.

Under additivity the yield on plot u with treatment k is

$$y_{uk} = x_u + t_k$$
$$= x_. + t_k + (x_u - x_.) ,$$

where x_u is the basal yield of the u-th unit, $x_.$ is the mean basal yield, and t_k is the yield of treatment k on plot u. The observed value of the f-th realization of treatment k can be expressed in terms of the set of conceptual values, y_{uk}, by

$$y_{kf} = \sum_{u=1}^{n} \delta_u^{kf} y_{uk}$$

$$= x_. + t_k + \sum_{u=1}^{n} \delta_u^{kf} e_u$$

where $e_u = (x_u - x_.)$. The error sum of squares, denoted by E.S.S., is given by

$$\text{E.S.S.} = \sum_{k,f} y_{kf}^2 - R \sum_{k} y_{k.}^2 .$$

We then find that

$$E(\text{E.S.S.}) = T(R-1)/(RT-1)S_2$$

8

and

$$V(E.S.S.) = \frac{2(R-1)(T-1)(R^2T^2 - 3RT+3)}{R(RT-1)^2(RT-2)(RT-3)} S_2^2 - \frac{2T(T-1)(R-1)}{(RT-1)(RT-2)(RT-3)} S_4$$

where $\quad S_2 = \sum_{u=1}^{n} e_u^2 \quad$ and $\quad S_4 = \sum_{u=1}^{n} e_u^4$.

B. The Error Sum of Squares under Constrained Randomization

By the definition of constrained sets of plans the expectation of the error sum of squares under constrained randomization is the same as the expectation of the error sum of squares under complete randomization. The δ_u^{kf} under constrained randomization take on the same values, i.e. 0 and 1, as under complete randomization, but their distributional properties are different. The probability that a certain combination of the random variables takes on a specified value depends on the values of u. For example, for a given constrained set we may have that

$$P\left[\delta_u^{kf} = 1 \;\middle|\; \delta_{u'}^{kf'} = 1\right] = 1 \quad \text{while} \quad P\left[\delta_{u''}^{kf} = 1 \;\middle|\; \delta_{u'''}^{kf'} = 1\right] = 0 ,$$

where $u \neq u' \neq u'' \neq u'''$ and $f \neq f'$. For this reason it appears very difficult to determine the variance of the error sum of squares in general under constrained randomization.

The notion of constrained randomization was developed to overcome some of the undesirable effects of systematic arrangements of the units. Since a linear relationship of basal yields to unit number constitutes perhaps the simplest example of this type, we will now consider this

9

situation in some detail. Because of origin and scale invariance of analysis of variance procedures we will let the basal yield of each unit equal the unit number.

The reason for examining the notion of constrained randomization is to develop some understanding of the properties it induces with reference to tests of significance, estimation, and sensitivity. It seems reasonable to expect that the constrained set with the smallest variance of the error sum of squares will be the most sensitive for detecting treatment differences.

We now give two examples of constrained sets of plans along with the distribution of their error sum of squares, where we assume that the basal yields are equal to their respective unit numbers.

Let $R = T = 4$ and $N = 5$. Yates (1936) gave the R.B.I.B.D. plan which leads to the constrained set of 5 plans given in Table 1.

TABLE 1

SET A OF PLANS AND THE DISTRIBUTION OF ITS
ERROR SUM OF SQUARES

Plan No.	Unit No.																E.S.S.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1	a	a	a	a	b	b	b	b	c	c	c	c	d	d	d	d	20.0
2	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	320.0
3	a	b	c	d	b	a	d	c	c	d	a	b	d	c	b	a	340.0
4	a	b	c	d	d	c	b	a	b	a	d	c	c	d	a	b	340.0
5	a	b	c	d	c	d	a	b	d	c	b	a	b	a	d	c	340.0

10

The error sum of squares for plan 5, for example, is

$$1^2 + 7^2 + 12^2 + 14^2 - (1 + 7 + 12 + 14)^2 / 4$$

$$+ 2^2 + 8^2 + 11^2 + 13^2 - (2 + 8 + 11 + 13)^2 / 4$$

$$+ 3^2 + 5^2 + 10^2 + 16^2 - (3 + 5 + 10 + 16)^2 / 4$$

$$+ 4^2 + 6^2 + 9^2 + 15^2 - (4 + 6 + 9 + 15)^2 / 4$$

$$= 340 .$$

For this constrained set

$$E(E.S.S.) = 272$$

and $\qquad V(E.S.S.) = 15,936 .$

Under complete randomization

$$E(E.S.S.) = 272$$

and $\qquad V(E.S.S.) = 2,339.20 .$

By permuting the rows and columns of the 16 x 16 square array for the constrained set given in Table, we find the constrained set of plans given in Table 2.

For this constrained set

$$E(E.S.S.) = 272$$

and $\qquad V(E.S.S.) = 38.40 .$

11

TABLE 2

SET B OF PLANS AND THE DISTRIBUTION OF ITS
ERROR SUM OF SQUARES

Plan No.	Unit No. 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	E.S.S.
1	a	b	b	a	b	a	c	d	c	d	d	c	d	c	a	b	276
2	a	a	b	b	c	c	d	d	a	a	b	b	c	c	d	d	260
3	a	b	c	d	a	b	b	a	d	c	b	a	d	c	c	d	276
4	a	b	c	d	d	c	d	c	c	d	a	b	b	a	b	a	276
5	a	b	a	b	c	d	a	b	c	d	c	d	a	b	c	d	272

The large numerical value of the variance of the error sum of squares for the constrained set of plans given in Table 1 was generated from a 4 x 4 completely orthogonalized Latin square. The above example is an illustration of the following general theorem.

Theorem 3:

If the unit number is equal to the basal yield of that unit and if a completely orthogonalized Latin square of side R exists, then in the class of all constrained sets with parameters $R = T$ and $N = R + 1$, the maximum value of the variance of the error sum of squares is attained by every constrained set corresponding to any completely orthogonalized Latin square of side R.

IV. POWER OF THE RANDOMIZATION TEST FOR CONSTRAINED RANDOMIZATION

A. Introduction

The common use of either the F distribution or the beta distribution, in connection with tests associated with the analysis of variance table, has its origin in the normality assumption about the distribution of the basic errors occurring in the models appropriate to the mathematical representation of basic experimental situations. Consideration of experiments where constrained randomization is useful, however, shows that the experimental unit errors follow a distribution distinctly non-normal.

We can in general perform an exact test as follows. Compute the value of the test criterion treatment sum of squares over total sum of squares, denoted by B_1, from the experiment as it was realized. Assuming that the values observed in the experiment are values characteristic of the experimental units proper, compute the values of the test criterion treatment sum of squares over total sum of squares for all potential realizations of the experiment. Arrange these computed values in ascending order of magnitude. Then on the basis of a significance level of say 5%, the null hypothesis of no treatment effect will be rejected when B_1 is among the upper 5% of the computed values in the ascending array. The above test procedure is exact in the sense that its significance level has the specified size. It is evident, however, that for the larger experiments, without using electronic computer methods, the computations required for performing the test are prohibitively laborious.

13

The performance of the experiment will yield a set of RT observed

yields and corresponding to these observed yields the value

B_1 = treatment sum of squares over total sum of squares can be

computed. Under the assumption that there are no treatment effects, any

of the other N assignments of the T treatments to the RT experi-

mental units will result in a set of RT observed yields. There are N

such sets of RT observed yields and for each set the value of B can be

computed. The N values so computed can now be arranged in an array

of ascending magnitudes. The position of B_1 in this array then tells us

the levels of significance for which the null hypothesis of no difference

among the T treatments would be rejected. In practice the actual

values of the test criterion treatment sum of squares over total sum of

squares need not be computed because there is a one-to-one corre-

spondence between the magnitudes of B and the magnitudes of

$$\sum_{j=1}^{T} T_j^2 \ ,$$

where T_j is the sum of the observed yields which receive treatment j,

so that an equivalent simplified test procedure can be specified.

To illustrate the idea, suppose that we used the constrained design

for 3 treatments on 6 units given on page 7. That is, we choose at

random one of the following 5 plans:

14

| Plan | Unit No. | | | | | |
No.	1	2	3	4	5	6
I	a	a	b	c	b	c
II	a	b	a	c	c	c
III	c	b	b	a	c	c
IV	a	b	c	b	a	c
V	a	b	c	c	b	a

We would of course, assign the letters a, b and c to the 3 treatments at random. Now suppose we drew plan IV, and obtained the following observations:

	Unit No.					
	1	2	3	4	5	6
Treatment	a	b	c	b	a	c
Response	15	18	21	20	19	24 .

To perform the randomization test we superimpose successively plans I, II, III, IV, and V on the observations, and analyze the resulting configurations. We obtain the following numbers:

Plan No.	Total S.Sq. around mean	Correction factor	Sum of (treatment total)^2	Treatment S.Sqs.	Error S.Sqs.	F
I	45.5	2281.5	2312.5	31.0	14.5	3.21
II	45.5	2281.5	2290.5	9.0	36.5	0.37
III	45.5	2281.5	2297.5	16.0	29.5	0.81
IV	45.5	2281.5	2322.5	41.0	4.5	13.67
V	45.5	2281.5	2285.5	4.0	41.5	0.14

For instance if we superimpose plan I on the observations, we obtain the configuration:

	Unit No.					
	1	2	3	4	5	6
Treatment	a	a	b	c	b	c
Response	15	18	21	20	19	24

The total sum of squares is

$$15^2 + 18^2 + 21^2 + 20^2 + 19^2 + 24^2 - \frac{(15 + 18 + 21 + 20 + 19 + 24)^2}{6} = 45.5 .$$

The correction factor is

$$\frac{(15 + 18 + 21 + 20 + 19 + 24)^2}{6} = 2,281.5 .$$

The sum of squares of treatment totals is

$$\frac{(15 + 18)^2}{2} + \frac{(21 + 19)^2}{2} + \frac{(20 + 24)^2}{2} = 2,312.5 .$$

The treatment sum of squares is

$$2,312.5 - 2,281.5 = 31.0 .$$

The error sum of squares is

$$45.5 - 31.0 = 14.5 .$$

16

The value for F is

$$\left(\frac{31.0}{2}\right) \div \left(\frac{14.5}{3}\right) \; = \; 3.21.$$

To make the randomization F test we note that the observed value of

F was 13.67 and that other possible plans we could have obtained

would give F values of 3.21, 0.37, 0.81, and 0.14. Therefore the

observed F value is equalled or exceeded in the totality of 5 plans

only once, and the significance level to be attached to the observed

differences is $\frac{1}{5}$ or 20%. Because the total sum of squares is constant,

there is a monotone relationship between F and the treatment sum of

squares, and because the correction factor is constant, one need consider

only the sum of squares of treatment totals in deciding the randomization

significance of the observed differences. Note in passing that the

assumption of normality would allow us to use the F table with 2 and

3 degrees of freedom leading to a significance level between 10% and

5% for the observed differences. In general the randomization test

leads to a discrete set of possible significance levels, in contrast to a

test based on an assumed continuous distribution.

The value of the power function for a particular vector of values,

$\delta_0 = (\delta_1, \delta_2, \ldots, \delta_T)$ of the treatments and a particular vector of values,

$P_0 = (\rho_1, \rho_2, \ldots, \rho_n)$, $n = RT$, of the units may be computed as follows.

Each arrangement of the T treatments on the RT experimental units

would yield a set of RT observed yields, y_i, for $i = 1, 2, \ldots, RT$. On

the basis of that particular set of observed yields, y_i, we may perform

the test of the hypothesis of no treatment differences as described in the

paragraph above, and hence record whether the hypothesis would be accepted or rejected for the set under consideration. Altogether N such sets of RT observed yields can be generated and for each set a decision as to the acceptance or rejection of the hypothesis can be made. The proportion of rejections of the null hypothesis at a particular significance level, constitutes the value of the power function of the test when the vector of treatment values is δ_0.

B. Empirical Results

Empirical results for the power of the randomization test under constrained randomization have been obtained for a number of different sets of values for the parameters R, T, and N. For each set of values R, T, and N, we consider three different constrained sets characterized by a small, an average, and a large value of the variance of the error sum of squares under linear basal yields. The different sets are denoted by L, M, and H respectively. For each such constrained set, L, M, and H, three different types of basal yields, x_i, are considered. We will refer to these types as

(1) Linear - these are nearly linearly distributed random variables with variance equal to K,

(2) Random - these are normally and independently distributed random variables with a mean of zero and a variance equal to K, and

(3) Semi-Random - these are distributed as the sum of a linearly distributed random variable with variance equal to

18

$\frac{1}{2}$ K and a normally and independently distributed random variable with a mean of zero and a variance equal to $\frac{1}{2}$ K.

The vectors of treatment effects are determined by the following relationships:

$$\sum_{j=1}^{T} t_j = 0$$

and

$$t_1 = a, \quad t_2 = a + b, \quad t_3 = a + 2b, \ldots, t_T = a + (T-1)b,$$

where a is arbitrary and b is chosen so that the ratio of the true standard deviation among the treatment effects to $\sqrt{K/R}$, which is the standard deviation of a treatment mean, equals λ, where λ takes on the values

$$0, \quad \frac{1}{12}, \quad \frac{1}{6}, \quad \frac{1}{4}, \quad \frac{1}{3}, \quad \frac{1}{2}, \quad \frac{2}{3}, \quad 1, \quad 1\frac{1}{3}, \quad 1\frac{2}{3}, \quad 2, \quad 3, \quad 4, \quad 5.$$

We now give some representative tables of the power function of the randomization test for the constrained sets considered.

TABLE 3

POWER OF THE RANDOMIZATION TEST AT A SIGNIFICANCE LEVEL OF 14.28% FOR THREE CONSTRAINED SETS WITH PARAMETERS R = 4, T = 2, AND N = 7

Power in Percent for Three Constrained Sets with

λ	Basal Yields Linear			Basal Yields Random			Basal Yields Semi-Random		
	L	M	H	L	M	H	L	M	H
0	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28
$\frac{1}{2}$	28.57	14.28	14.28	21.42	14.28	14.28	28.57	21.42	14.28
$\frac{2}{3}$	42.85	14.28	14.28	35.71	28.57	14.28	28.57	28.57	14.28
1	50.00	21.42	14.28	42.85	42.85	14.28	50.00	28.57	14.28
$1\frac{1}{3}$	57.14	50.00	28.57	57.14	50.00	28.57	50.00	57.14	14.28
$1\frac{2}{3}$	71.42	50.00	50.00	64.28	71.42	50.00	71.42	71.42	14.28
2	85.71	78.57	71.42	85.71	85.71	78.57	71.42	78.57	42.85
3	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
V(E.S.S.)	23	63	114	23	40	92	30	65	276

TABLE 4

POWER OF THE RANDOMIZATION TEST AT A SIGNIFICANCE LEVEL OF 14.28% FOR THREE CONSTRAINED SETS WITH PARAMETERS R = 3, T = 5, AND N = 7

Power in Percent for Three Constrained Sets with

λ	Basal Yields Linear			Basal Yields Random			Basal Yields Semi-Random		
	L	M	H	L	M	H	L	M	H
0	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28
$\frac{1}{6}$	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28
$\frac{1}{3}$	24.52	14.28	14.28	14.28	14.16	14.28	14.28	14.28	18.92
$\frac{2}{3}$	45.35	17.02	14.28	15.95	23.21	18.33	15.47	14.76	28.09
$1\frac{2}{3}$	87.73	84.04	74.40	51.54	59.52	66.66	64.40	62.02	75.83
2	96.42	96.78	97.97	73.09	79.88	79.64	84.52	86.66	89.16
3	100.00	100.00	100.00	100.00	99.88	98.92	100.00	100.00	99.88
4	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
5	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
V(E.S.S.)	438	2896	13356	11450	8230	5307	6630	6263	2175

TABLE 5

POWER OF THE RANDOMIZATION TEST AT A SIGNIFICANCE LEVEL OF 20% FOR THREE CONSTRAINED SETS WITH PARAMETERS R = 4, T = 4, AND N = 5

Power in Percent for Three Constrained Sets with

λ	Basal Yields Linear			Basal Yields Random			Basal Yields Semi-Random		
	L	M	H	L	M	H	L	M	H
0	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00
$\frac{1}{4}$	43.33	20.00	20.00	20.83	20.00	20.83	20.00	20.00	20.00
$\frac{1}{2}$	62.50	20.00	20.00	25.83	21.66	28.33	21.66	20.00	23.22
$\frac{2}{3}$	68.33	22.50	20.00	38.33	30.00	45.00	27.50	24.16	25.83
1	70.83	32.50	20.00	64.16	60.00	68.33	46.66	44.16	30.00
$1\frac{1}{3}$	82.50	54.16	20.00	87.50	81.66	85.83	76.66	74.16	71.66
$1\frac{2}{3}$	86.66	75.83	22.50	97.50	97.50	97.50	91.66	90.00	94.16
2	99.16	89.16	28.33	100.00	100.00	100.00	99.16	99.16	95.83
3	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
V(E.S.S.)	38	3179	15936	670	834	322	1484	1559	2452

TABLE 6

POWER OF THE RANDOMIZATION TEST AT A SIGNIFICANCE LEVEL OF 9.09% FOR THREE CONSTRAINED SETS WITH PARAMETERS R = 3, T = 4, AND N = 11

Power in Percent for Three Constrained Sets with

λ	Basal Yields Linear			Basal Yields Random			Basal Yields Semi-Random		
	L	M	H	L	M	H	L	M	H
0	9.09	9.09	9.09	9.09	9.09	9.09	9.09	9.09	9.09
$\frac{1}{4}$	13.25	10.22	9.09	9.09	9.09	8.33	15.15	7.95	9.09
$\frac{1}{2}$	20.83	11.74	9.09	9.09	11.74	10.60	23.10	15.15	9.09
$\frac{2}{3}$	22.34	16.28	9.09	14.01	15.15	12.87	25.75	21.59	9.09
1	28.41	23.10	9.09	28.03	32.95	27.65	41.28	37.87	13.63
$1\frac{1}{3}$	41.66	36.36	14.01	40.15	49.62	43.18	58.33	57.95	48.86
$1\frac{2}{3}$	54.54	51.89	33.71	59.47	63.63	59.84	79.16	76.89	81.44
2	69.69	66.28	63.63	74.24	80.68	79.16	92.42	89.39	93.88
5	100.00	100.00	100.00	100.00	99.62	100.00	100.00	99.62	100.00
V(E.S.S.)	444	593	1372	930	1064	1096	388	545	1129

TABLE 7

POWER OF THE RANDOMIZATION TEST AT A SIGNIFICANCE LEVEL OF 9.09% FOR THREE CONSTRAINED SETS WITH PARAMETERS R = 4, T = 3, AND N = 11

Power in Percent for Three Constrained Sets with

λ	Basal Yields Linear			Basal Yields Random			Basal Yields Semi-Random		
	L	M	H	L	M	H	L	M	H
0	9.09	9.09	9.09	9.09	9.09	9.09	9.09	9.09	9.09
$\frac{1}{3}$	10.60	9.09	9.09	9.09	9.09	13.63	9.09	15.15	16.66
$\frac{1}{2}$	12.12	10.60	9.09	9.09	10.60	19.69	9.09	27.27	19.69
$\frac{2}{3}$	16.66	9.09	9.09	10.60	13.63	25.75	9.09	34.84	21.21
1	22.72	16.66	9.09	19.69	27.27	33.33	30.30	45.45	45.45
$1\frac{1}{3}$	40.91	34.84	9.09	34.84	37.87	48.48	56.06	63.63	74.24
$1\frac{2}{3}$	53.03	53.03	21.21	50.00	56.06	66.66	78.78	80.30	86.36
2	72.72	66.66	56.06	77.27	75.75	80.30	92.42	92.42	96.97
4	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
V(E.S.S.)	465	555	1076	409	386	268	158	68	36

24

TABLE 8

POWER OF THE RANDOMIZATION TEST AT A SIGNIFICANCE LEVEL OF 14.28% FOR THREE CONSTRAINED SETS WITH PARAMETERS R = 2, T = 4, AND N = 7

Power in Percent for Three Constrained Sets with

λ	Basal Yields Linear			Basal Yields Random			Basal Yields Semi-Random		
	L	M	H	L	M	H	L	M	H
0	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28	14.28
$\frac{1}{3}$	19.04	17.26	14.28	20.83	19.64	17.85	14.28	14.28	14.28
$\frac{1}{2}$	22.62	19.64	14.28	24.40	21.42	23.81	14.28	14.28	14.28
$\frac{2}{3}$	26.19	22.02	14.28	32.73	24.40	28.38	14.28	15.47	14.88
1	35.12	29.76	19.04	43.45	27.38	42.26	16.07	16.66	16.66
$1\frac{1}{3}$	40.47	41.07	35.12	60.71	36.31	61.90	18.45	19.64	23.81
$1\frac{2}{3}$	58.92	52.97	52.38	77.38	54.16	77.97	25.00	27.97	31.54
2	73.21	64.88	71.42	92.85	79.76	95.23	32.14	34.52	35.12
3	95.83	97.02	97.62	100.00	95.23	100.00	61.90	63.09	60.71
4	100.00	100.00	100.00	100.00	100.00	100.00	83.33	81.54	83.33
V(E.S.S.)	51	116	126	20	36	18	700	690	758

25

C. Discussion of the Empirical Results

Consider the three different constrained sets with the same values for the parameters R, T, and N, for which the planned distinguishing feature consists in their different values for the variance of the error sum of squares. Under nearly linear basal yields the power of the randomization test for the constrained set with a low variance of the error sum of squares is nearly always greater than or equal to that for either of the constrained sets with an average or a high variance of the error sum of squares. Similar differences in the power occur for the constrained sets with average and high variances of the error sum of squares. The above is also true, but somewhat less pronounced, for the semi-randomly distributed basal yields. Also, when the basal yields form a sample from a normal distribution, the differences between the power functions of the three different constrained sets are small. As an example of these facts we refer to Table 5. When $\lambda = 1\frac{1}{3}$ and the basal yields are nearly linear the power is 82.50%, 54.16%, and 20.00% for the constrained sets L, M, and H, respectively. With the same situation for random basal yields the powers are 87.50%, 81.66%, and 85.83% and 76.66%, 74.16%, and 71.66% for semi-random basal yields. It should be noted that the variances of the error sum of squares for all three plans with random and semi-random basal yields are all small relative to the two large values of the variance of the error sum of squares under nearly linear basal yields.

One would expect that a constrained set, for which the variance of the error sum of squares is nearly equal to the variance of the error

sum of squares under complete randomization, has a power function nearly equal to that holding under complete randomization. Thus, with the above assumption the tables indicate that with nearly linear basal yields a constrained set, for which the variance of the error sum of squares is less than that under complete randomization, has greater power than that holding under complete randomization. Table 7 shows that increases in λ are not always accompanied by corresponding increases in the power. This last fact does not correspond to the normal theory result where the power function of the analysis of variance test is a strictly increasing function of λ.

V. LITERATURE CITED

Anscombe, F. J. 1948. The validity of comparative experiments. Journal of the Royal Statistical Society Ser. A, 111: 181-197.

Ball, Rouse W. W. 1937. Mathematical recreations and essays. 10th ed. MacMillan Company, Limited, London.

Bhattacharya, K. N. 1945. The B.I.B.D. $v = b = 31$, $r = k = 10$, $f = 3$. Proceedings of the Indian Science Congress, 32 I.S.C.: 166. Original not available; cited in Shrikhande, S.S. 1951. On the non-existence of affine resolvable balanced incomplete block designs. Sankhyā, 11: 186.

Bose, R. C. 1939. On the construction of balanced incomplete block designs. Annals of Eugenics, 9: 353-398.

_____ 1942a. On some new series of balanced incomplete block designs. Bulletin of the Calcutta Mathematical Society, 34: 17-31.

_____ 1942b. A note on two series of balanced incomplete block designs. Bulletin of the Calcutta Mathematical Society, 34: 129-130.

_____ 1942c. A note on the resolvability of balanced incomplete block designs. Sankhyā, 6: 105-110.

_____ 1947. On a resolvable series of balanced incomplete block designs. Sankhyā, 8: 249-256.

_____, S. S. Shrikhande and K. N. Bhattacharya. 1953. On the construction of group divisible incomplete block designs. Annals of Mathematical Statistics, 24: 167-195.

Carmichael, R. D. 1937. Introduction to the theory of groups of finite order. Ginn and Company, Boston.

Cox, D. R. 1956. A note on weighted randomization. Annals of Mathematical Statistics, 27: 1144-1150.

Cox, G. M. 1940. Enumeration and construction of balanced incomplete block configurations. Annals of Mathematical Statistics, 11: 72-85.

Das, M. N. 1954. On parametric relations in a balanced incomplete block design. Journal of the Indian Society of Agricultural Statistics, 6: 147-152.

Fisher, R. A. 1940. An examination of the different possible solutions of a problem in incomplete blocks. Annals of Eugenics, 10: 52-75.

Grundy, P. M. and M. J. R. Healy. 1950. Restricted randomization and quasi-Latin squares. Journal of the Royal Statistical Society Ser. B, 12: 286-291.

Kempthorne, O. 1952. The design and analysis of experiments. John Wiley and Sons, Incorporated, New York.

_____ 1955. The randomization theory of experimental inference. Journal of the American Statistical Association, 50: 946-967.

Kishen, K. and C. R. Rao. 1952. An examination of various inequality relations among parameters of the balanced incomplete block design. Journal of the Indian Society of Agricultural Statistics, 4: 137-144.

Majumdar, K. N. 1953. On some theorems in combinatorics relating to B.I.B. designs. Annals of Mathematical Statistics, 24: 377-389.

Mikhail, W. F. 1960. An inequality for balanced incomplete block designs. Annals of Mathematical Statistics, 31: 520-522.

Nair, K. R. 1943. Certain inequality relationships among the combinatorial parameters of incomplete block designs. Sankhyā, 6: 255-259.

Plackett, R. L. 1960. Principles of regression analysis. Oxford University Press, London.

Primrose, E. J. F. 1952. Resolvable balanced incomplete block designs. Sankhyā, 12: 137-140.

Rao, C. R. 1961. A study of B.I.B. designs with replications 11 to 15. Sankhyā, 23: 117-128.

Roy, P. M. 1952. A note on the resolvability of B.I.B.D. Bulletin of the Calcutta Statistical Association, 4: 130-133.

Schutenberger, M. P. 1949. A non-existence theorem for an infinite family of symmetrical block designs. Annals of Eugenics, 14: 286-287.

Shrikhande, S. S. 1951. On the non-existence of affine resolvable balanced incomplete block designs. Sankhyā, 11: 185-186.

_____ 1953. The non-existence of certain affine resolvable balanced incomplete block designs. Canadian Journal of Mathematics, 5: 413-420.

_____ 1954. Affine resolvable balanced incomplete block designs and non-singular group divisible designs. Bulletin of the Calcutta Statistical Association, 5: 139-142.

Sprott, D. A. 1956. Some series of balanced incomplete block designs. Sankhyā, 17: 185-193.

Sutter, G. J. 1962. Some aspects of constrained randomization. Unpublished M.S. Thesis, Iowa State University of Science and Technology, Ames, Iowa.

Wilk, M. B. 1955. The randomization analysis of a generalized randomized block design. Biometrika, 42: 70-79.

Yates, F. 1936. Incomplete randomized blocks. Annals of Eugenics, 7: 121-140.

Youden, W. J. 1956. Randomization and experimentation. Annals of Mathematical Statistics, 27: 1185-1186.

Aeronautical Research Laboratories,Wright-
Patterson AFB, Ohio.   SOME ASPECTS OF CON-
STRAINED RANDOMIZATION, by Gary J. Sutter,
George Zyskind, Oscar Kempthorne, Iowa State
University, Ames, Iowa. November 1962. 30p.
incl. tables (Project 7071; Task 7071-01).
(Contract AF33(616)-8269) (ARL

Unclassified Report

   In Chapter I the definition of con-
strained sets of plans and the problems to
be considered are given. In Chapter II a
correspondence between constrained sets of
plans and RT x RT square arrays and a corre-
spondence between constrained sets of plans
and resolvable balanced incomplete block

(over)

designs are given. These correspondences
led to the enumeration of many constrained
designs. In Chapter II some relationships
between the variance of the error sum of
squares under complete randomization and
under constrained randomization are given.
In Chapter IV the power of the randomization
test under constrained randomization was
considered for nearly linearly, randomly,
and semi-randomly distributed basal yields.
The empirical results exhibit a tendency of
the sensitivity of the test for the detection
of treatment differences to increase with
decreases in the variance of the error sum
of squares of the constrained sets.