STARTED Monday 4/11/22 (Week 12, Lecture 31)
@45min mark

HANDOUT # 10

FRACTIONAL FACTORIAL DESIGNS (FF)

I. Example Illustrating the Need for FF

II. FF - 2^{n-p} Designs

a. Contrasts to Test for Effects in a 2^n Design

b. Confounded Effects

c. Resolution of a 2^{n-p} Design

d. Constructing a 2^{n-p} Design

e. Aliasing Sets

f. Selecting the Appropriate 2^{n-p} Design

g. Minitab Output for Designing a 2^{n-p} Design

h. SAS Output for Designing a 2^{n-p} Design

III. Screening Designs and Sequential Experimentation

a. Plackett-Burman Designs

b. Example

IV. Analysis of Data From a FF Design

V. 3^n Designs and Mixed Level Designs

- Supplemental Reading: Design & ANOVA Book - Chapter 15

I. Fractional Factorial Experiments

With Selected Material from *Statistical Design and Analysis of Experiments*

by Robert Mason, Richard Gunst, and James Hess

Fractional factorial experiments are important alternatives to complete factorial experiments when cost, time, or experimental constraints preclude the execution of complete factorial experiments. In addition, experiments that involve many factors are routinely conducted as fractional factorials because it is not necessary to test all possible factor-level combinations in order to estimate the most crucial factor effects, main effects and low-order interactions.

An experiment was designed to study the corrosion rate of a reactor at a chemical acid plant. There were five factors of interest: each having two levels.

Factor Levels for Acid-Plant Corrosion-Rate Study

Factor	Levels
A_1: Raw-material feed rate	3000, 6000 pph
A_2: Gas temperature	$100°C$, $220\ °C$
A_3: Scrubber water	5%, 20%
A_4: Reactor-bed acid	20%, 30%
A_5: Exit temperature	$300°C$, $360°C$

Initially it was unknown whether changing any of these five process variables (factors) would have an effect on the reactor corrosion rate, the response variable of interest to the process engineers. In order to minimize maintenance and downtime, it was desirable to operate this acid plant under process conditions which produce a low corrosion rate. If one were to test all possible combinations of three or four levels of the five process variables, A_1, A_2, A_3, A_4, A_5:

$$\underline{\ 3\ }\ \underline{\ 3\ }\ \underline{\ 3\ }\ \underline{\ 3\ }\ \underline{\ 3\ } = 3^5 = 243 \text{ runs or } \underline{\ 4\ }\ \underline{\ 4\ }\ \underline{\ 4\ }\ \underline{\ 4\ }\ \underline{\ 4\ } = 4^5 = 1024 \text{ runs}$$

* we'll focus on mainly 2 L designs.

would be needed for a single replication of the possible treatments. A practical difficulty with this investigation was that the plant needed to cease commercial production for the duration of the study. An experimental design that required 243 or 1024 runs would have been prohibitively expensive. A small screening experiment was designed in which only gross factor effects were identified, the feasible extremes of the five variables. It is still necessary to execute $2^5 = 32$ runs. However, the cost of experimentation would be greatly reduced over an experiment involving three or four levels of each of the factors.

The table on the next page displays this proposed design using the notation: -1 for the low level of the factor and $+1$ for the high level of the factor.

An alternative notation designates the 32 treatments: If a factor is at its low level then its symbol is not displayed. Thus, we have that Treatment $a_1a_2a_5$ represents Factors

A_1, A_2, A_5 at their High Level; and Factors A_3, A_4 at their Low Level

The treatment having all Factors at their low level is designated as I or as (1).

The following table illustrates this notation for all $2^5 = 32$ treatments:

Table 1: Treatments 2^5 Factorial

TRT	A_1	A_2	A_3	A_4	A_5	TREATMENT	Y_{ijklm}
1	-1	-1	-1	-1	-1	(I)	Y_{11111}
2	-1	-1	-1	-1	1	a_5	Y_{11112}
3	-1	-1	-1	1	-1	a_4	Y_{11121}
4	-1	-1	-1	1	1	a_4a_5	Y_{11122}
5	-1	-1	1	-1	-1	a_3	Y_{11211}
6	-1	-1	1	-1	1	a_3a_5	Y_{11212}
7	-1	-1	1	1	-1	a_3a_4	Y_{11221}
8	-1	-1	1	1	1	$a_3a_4a_5$	Y_{11222}
9	-1	1	-1	-1	-1	a_2	Y_{12111}
10	-1	1	-1	-1	1	a_2a_5	Y_{12112}
11	-1	1	-1	1	-1	a_2a_4	Y_{12121}
12	-1	1	-1	1	1	$a_2a_4a_5$	Y_{12122}
13	-1	1	1	-1	-1	a_2a_3	Y_{12211}
14	-1	1	1	-1	1	$a_2a_3a_5$	Y_{12212}
15	-1	1	1	1	-1	$a_2a_3a_4$	Y_{12221}
16	-1	1	1	1	1	$a_2a_3a_4a_5$	Y_{12222}
17	1	-1	-1	-1	-1	a_1	Y_{21111}
18	1	-1	-1	-1	1	a_1a_5	Y_{21112}
19	1	-1	-1	1	-1	a_1a_4	Y_{21121}
20	1	-1	-1	1	1	$a_1a_4a_5$	Y_{21122}
21	1	-1	1	-1	-1	a_1a_3	Y_{21211}
22	1	-1	1	-1	1	$a_1a_3a_5$	Y_{21212}
23	1	-1	1	1	-1	$a_1a_3a_4$	Y_{21221}
24	1	-1	1	1	1	$a_1a_3a_4a_5$	Y_{21222}
25	1	1	-1	-1	-1	a_1a_2	Y_{22111}
26	1	1	-1	-1	1	$a_1a_2a_5$	Y_{22112}
27	1	1	-1	1	-1	$a_1a_2a_4$	Y_{22121}
28	1	1	-1	1	1	$a_1a_2a_4a_5$	Y_{22122}
29	1	1	1	-1	-1	$a_1a_2a_3$	Y_{22211}
30	1	1	1	-1	1	$a_1a_2a_3a_5$	Y_{22212}
31	1	1	1	1	-1	$a_1a_2a_3a_4$	Y_{22221}
32	1	1	1	1	1	$a_1a_2a_3a_4a_5$	Y_{22222}

Half (−) half (+)

for each group of (−) or (+) for the previous ~~factor~~, half of these are (−) half (+)

same thing is repeated for k>3.

STOP Monday 4/11/22 (Week 12, Lecture 31)

START Wednesday 4/13/22 (week 12, lecture 35)

The table of treatments on the previous page will be used to produce estimates of the various effects. The fact that each of the factors has only two levels, L and H, greatly simplifies the calculations.

Main Effects: Factor A_1 has only two levels hence the estimated

main effect of Factor A_1 is $\hat{\mu}_{2....} - \hat{\mu}_{1....} = \bar{Y}_{2....} - \bar{Y}_{1....}$

This quantity can be obtained by multiply the transpose of the column vector, $\mathbf{A}_1$ by the column of response values $\mathbf{Y}$ times 2/n: That is,

Main Effect of Factor A_1 :

$$\begin{aligned}
\frac{2}{32}\,\mathbf{A}_1^{'} \times \mathbf{Y} &= (-Y_{11111} - Y_{11112} - Y_{11121} - Y_{11122} - Y_{11211} - Y_{11212} - Y_{11221} - Y_{11222} \\
&\quad -Y_{12111} - Y_{12112} - Y_{12121} - Y_{12122} - Y_{12211} - Y_{12212} - Y_{12221} - Y_{12222})/16 \\
&\quad +(Y_{21111} + Y_{21112} + Y_{21121} + Y_{21122} + Y_{21211} + Y_{21212} + Y_{21221} + Y_{21222} \\
&\quad +Y_{22111} + Y_{22112} + Y_{22121} + Y_{22122} + Y_{22211} + Y_{22212} + Y_{22221} + Y_{22222})/16 \\
&= -\bar{Y}_{1....} + \bar{Y}_{2....}
\end{aligned}$$

Similarly, the estimated main effect of factor A_4 is $\hat{\mu}_{...2.} - \hat{\mu}_{...1.} =$

$$\text{Main Effect of Factor } A_4 : \frac{1}{16}\,\mathbf{A}_4^{'} \times \mathbf{Y} = -\bar{Y}_{...1.} + \bar{Y}_{...2.}$$

Two-way Interactions: The estimated two-way interaction between Factors A_1 and A_2 is given by

$$(\bar{Y}_{11...} - \bar{Y}_{12...}) - (\bar{Y}_{21...} - \bar{Y}_{22...}) = \bar{Y}_{11...} - \bar{Y}_{12...} - \bar{Y}_{21...} + \bar{Y}_{22...}$$

This expression can be obtained by first obtaining the entrywise product matrix $\mathbf{A}_1 \circ \mathbf{A}_2$ which is obtained by multiplying the elements of $\mathbf{A}_1$ by the corresponding coefficients of $\mathbf{A}_2$ to form a $n \times 1$ vector. The two-way interaction between Factors A_1 and A_2 is then given by

$$\frac{4}{n}\,(\mathbf{A}_1 \circ \mathbf{A}_2)^{'} \times \mathbf{Y}$$

In our example, we would have that the two way interaction is given by

$$\begin{aligned}
\frac{1}{8}\,(\mathbf{A}_1 \circ \mathbf{A}_2)^{'} \times \mathbf{Y} &= (Y_{11111} + Y_{11112} + Y_{11121} + Y_{11122} + Y_{11211} + Y_{11212} + Y_{11221} + Y_{11222})/8 \\
&\quad +(-Y_{12111} - Y_{12112} - Y_{12121} - Y_{12122} - Y_{12211} - Y_{12212} - Y_{12221} - Y_{12222})/8 \\
&\quad +(-Y_{21111} - Y_{21112} - Y_{21121} - Y_{21122} - Y_{21211} - Y_{21212} - Y_{21221} - Y_{21222})/8 \\
&\quad +(Y_{22111} + Y_{22112} + Y_{22121} + Y_{22122} + Y_{22211} + Y_{22212} + Y_{22221} + Y_{22222})/8 \\
&= (\bar{Y}_{11...} - \bar{Y}_{12...}) - (\bar{Y}_{21...} - \bar{Y}_{22...})
\end{aligned}$$

The estimated 3-way and higher interactions are obtained in a similar fashion. For example, the three-way interaction between factors A_1, A_2, A_3 is obtained from

$$\frac{8}{n}\,(\mathbf{A}_1 \circ \mathbf{A}_2 \circ \mathbf{A}_3)^{'} \times \mathbf{Y}$$

After proposing this experiment the statistician was told that the cost of 32 runs was too great and hence a design having only 16 runs was constructed. This represents $\frac{1}{2}$ of the 32 runs needed for a single replication of the 32 treatments. This type of a design is named a **fractional factorial design** since only a fraction of all possible treatment combinations are included. A typical designation of the design is given as

$(\frac{1}{2})^p(2^n) = 2^{n-p}$ fractional factorial design.

This notation identifies there are n factors each at two levels but only a 2^{-p} fraction of the 2^n possible treatments are included in the experiment. Thus, if we used 16 of the 32 possible treatments we would have only $\frac{1}{2} = 2^{-1}$ of all the treatments; hence a $\frac{1}{2}2^5 = 2^{5-1}$ fractional factorial design. If we used 8 of the 32 possible treatments we would have only $\frac{1}{4} = 2^{-2}$ of all the treatments; hence a $\frac{1}{4}2^5 = 2^{-2}2^5 = 2^{5-2}$ fractional factorial design.

Whenever fractional factorial experiments are conducted, some effects are confounded with one another.

> **Confounded effects.** Two or more experimental effects are confounded if calculated effects can only be attributed to their combined influence on the response, not to their individual ones. Two or more effects are confounded if the calculation of one effect uses the same (apart from sign) difference or contrast of the response averages as the calculation of the other effects.

According to this definition, one effect is confounded with another effect if the effects representations of the two effects are identical, apart from a possible sign reversal. Effects that are confounded in this way are called **aliases.**

Suppose we had five factors each at two levels but we able to only conduct 8 runs of the experiment. We would then take the 32 possible treatments and block them into 4 blocks of 8 treatments each. We would then have an incomplete block design. If possible, we would run the 4 blocks separately and then have an experiment under which we observed all 32 treatments. The following table illustrates the assignment of the treatments to the 4 blocks:

2-way interactions in Block 1

each column is a contrast for its respective factor.

not a contrast (elements must sum to 0)

confounded w/ A1

Table 2: Trt Effects 2^5 Factorial in 4 Blocks

BLOCK	TRT	A_1	A_2	A_3	A_4	A_5	$A_1A_2A_4$	$A_1A_3A_5$	$A_2A_3A_4A_5$	$A_1A_2A_3A_4A_5$
1	1	-1	-1	-1	-1	-1	-1	-1	1	-1
1	6	-1	-1	1	-1	1	-1	-1	1	-1
1	11	-1	1	-1	1	-1	-1	-1	1	-1
1	16	-1	1	1	1	1	-1	-1	1	-1
1	20	1	-1	-1	1	1	-1	-1	1	1
1	23	1	-1	1	1	-1	-1	-1	1	1
1	26	1	1	-1	-1	1	-1	-1	1	1
1	29	1	1	1	-1	-1	-1	-1	1	1
2	4	-1	-1	-1	1	1	1	1	1	-1
2	7	-1	-1	1	1	-1	1	1	1	-1
2	10	-1	1	-1	-1	1	1	1	1	-1
2	13	-1	1	1	-1	-1	1	1	1	-1
2	17	1	-1	-1	-1	-1	1	1	1	1
2	22	1	-1	1	-1	1	1	1	1	1
2	27	1	1	-1	1	-1	1	1	1	1
2	32	1	1	1	1	1	1	1	1	1
3	2	-1	-1	-1	-1	1	-1	1	-1	1
3	5	-1	-1	1	-1	-1	-1	1	-1	1
3	12	-1	1	-1	1	1	-1	1	-1	1
3	15	-1	1	1	1	-1	-1	1	-1	1
3	19	1	-1	-1	1	-1	-1	1	-1	-1
3	24	1	-1	1	1	1	-1	1	-1	-1
3	25	1	1	-1	-1	-1	-1	1	-1	-1
3	30	1	1	1	-1	1	-1	1	-1	-1
4	3	-1	-1	-1	1	-1	1	-1	-1	1
4	8	-1	-1	1	1	1	1	-1	-1	1
4	9	-1	1	-1	-1	-1	1	-1	-1	1
4	14	-1	1	1	-1	1	1	-1	-1	1
4	18	1	-1	-1	-1	1	1	-1	-1	-1
4	21	1	-1	1	-1	-1	1	-1	-1	-1
4	28	1	1	-1	1	1	1	-1	-1	-1
4	31	1	1	1	1	-1	1	-1	-1	-1

These are the trts. used in table 3.

P=I, def. cont. = $A_1A_2A_3A_4A_5$

Q: is A_2A_4 confounded w/ A_1? in block 1?. Similarly for A_5 & A_1A_5?

If we decided to run only 8 treatments, a 2^{5-2} fractional factorial ($8 = \frac{1}{4}32 = 2^{-2}2^5$) and used the treatments given in Block 1 in the above table, then the effects of the 3-way interactions $A_1A_2A_4$, $A_1A_3A_5$, and the 4-way interaction $A_2A_3A_4A_5$ would be non-estimable because their coefficients do not add to 0 and hence are not contrasts. Also, the main effect of A_1 would be confounded with the 5-way interaction $A_1A_2A_3A_4A_5$. However, note that none of the main effects would be confounded with each other since the pattern of -1's and $+1$'s is unique for each of the main effects.

The goal in the design of fractional factorial experiments is to ensure that the effects of primary interest are either not confounded with other effects or, if that is not possible, confounded with effects that are not likely to have appreciable magnitudes. Confounding was defined as the situation where an effect cannot unambiguously be attributed to a single main effect or interaction. We have previously shown that factor effects are defined in terms of contrasts of response averages. In a factorial design involving factors having only two levels, all effects are identified by contrasts involving only -1's and $+1$'s as was demonstrated for the 2^5 design given in Table 2.

For example, consider the 2^{5-2} fractional factorial given by the 8 treatments in Block 1. Let Y_i be the observation from treatment i. We have observations on Treatments $i = 1, 6, 11, 16$ (Factor A_1 at low level) and Treatments $i = 20, 23, 26, 29$ (Factor A_1 at high level). Thus, the main effect of factor A_1 would be estimated by taking the mean of the observations at the high level minus the mean of the observations at the low level of

$$A_1\colon\ M(A_1) = -(Y_1 + Y_6 + Y_{11} + Y_{16})/4 + (Y_{20} + Y_{23} + Y_{26} + Y_{29})/4$$

This is equivalent to using the contrast:

$$M(A_1) = -Y_1 - Y_6 - Y_{11} - Y_{16} + Y_{20} + Y_{23} + Y_{26} + Y_{29}.$$

Similarly, the main effect of factor A_2 is estimated by taking the mean of the observations at the high level minus the mean of the observations at the low level of A_2:

$$A_2\colon\ M(A_2) = -(Y_1 + Y_6 + Y_{20} + Y_{23})/4 + (Y_{11} + Y_{16} + Y_{26} + Y_{29})/4.$$

This yields the equivalent contrast:

$$M(A_2) = -Y_1 - Y_6 + Y_{11} + Y_{16} - Y_{20} - Y_{23} + Y_{26} + Y_{29}.$$

We then obtain the estimated contrast for the 2-way interaction between factors A_1 and A_2 by just multiplying the corresponding coefficients in the contrast for the two main effects yielding:

$$A_1 * A_2\colon\ I(A_1 * A_2) = +Y_1 + Y_6 - Y_{11} - Y_{16} - Y_{20} - Y_{23} + Y_{26} + Y_{29}.$$

If only the treatments given in Block 1 were used in the experiment, then we can easily determine using Table 2 which effects are confounded.

In general, when designing fractional factorial experiments one seeks to confound either effects known to be negligible relative to the uncontrolled experimental error variation, or in the absence of such knowledge, high-order interactions, usually those involving three or more factors. Confounding of high-order interactions is recommended because frequently these interactions either do not exist or are negligible relative to main effects and low-order interactions.

Confounding occur when a complete factorial experiment is conducted in blocks having only a portion of all the treatments in each block. Confounding also occurs when only a portion of all the possible factor-level combinations are included in a CRD design. Confounding occurs because two or more effects representations (apart from a change in all the signs) are the same. A calculated effect then represents the combined influence of the effects. In some instances (e.g., one-factor-at-a-time experiments) the confounding pattern may be so complex that one cannot state with assurance that any of the calculated effects measure the desired factor effects. This again is why planned confounding, confounding in which important effects either are not confounded or are only confounded with effects that are believed to be negligible, is the basis for the statistical constructions of fractional factorial experiments.

Design Resolution. An experimental design is of resolution **R** if all effects containing **s** or fewer factors are not confounded with any effects containing fewer than **R- s** factors.

Main Effects have **s=1**; two-way interactions have **s=2**; three-way interactions have **s=3**; etc

Design resolution is defined in terms of which effects are not confounded with which other effects. One ordinarily seeks fractional factorial experiments that have main effects (s = 1) and low-order interactions (say those involving s = 2 or 3 factors) not confounded with other main effects and low-order interactions-equivalently, in which they are confounded only with high-order interactions. The resolution of a design is usually denoted by capital Roman letters; e.g., III, IV, V. The symbol R in the above definition denotes the Arabic equivalent of the Roman numeral; e.g., R = 3,4, 5.

Resolution III Design: A design in which main effects are not confounded with other main effects but are confounded with two-factor and higher-order interactions. (also denoted R_{III} design.

A design which is of resolution III has all effects containing s = 1 factors (main effects) not confounded with effects containing fewer than R-s = 3-1 = 2 factors, that is, all effects having only 1 factor (main effects). That is, main effects (s=1) are not confounded with any other main effect. However, some of the main effects may be confounded with two-way or higher order interactions.

Resolution IV Design: A design in which main effects are not confounded with other main effects and two-factor interactions but some two-factor interactions are confounded with other two-factor interactions (also denoted R_{IV} design.

In a resolution-IV design, main effects (s=1) are not confounded with effects involving fewer than R-s = 4 -1 = 3 factors. Thus, in a R=IV design all main effects (s=1) are not confounded with any other main effect ($s = 1 < 3$) nor with any two-factor interactions ($s = 2 < 3$).

Two-factor interactions (s = 2) are not confounded with effects involving fewer than R-s = 4-2 = 2 factors, that is, with any main effects (s = 1). There are some two-factor interactions confounded with other two-factor or higher order interactions in R_{IV} designs.

Resolution V Design: A design in which main effects and two-factor interactions are not confounded with one another. Some main effects are confounded with four-factor and higher interactions; some two-factor interactions are confounded with three-factor and higher-order interactions.

A design of resolution V has main effects and two-factor interactions ($s = 1,2$) not confounded with one another. This results from having $R-1=5-1=4$ which implies that main effects ($s=1$) are not confounded with any other main effect $(s = 1 < 4)$ or two-way interaction ($s = 2 < 4$) or three-way interactions $(s = 3 < 4)$. However, some main effects are confounded with four factor and higher-order interactions.

Two factor interactions ($s=2$) are not confounded with any other two factor interaction $(s = 2 < 3 = R - 2)$ but are confounded with some three factor interactions $(s = 3)$ or higher order interactions.

Three factor interactions ($s=3$) are not confounded with any main effect $(s = 1 < 2 = R - 3)$ but are confounded with some two factor or higher order interactions.

A design of resolution VI has main effects and two factor interactions not confounded with one another. Some main effects are confounded with five factor and higher order interactions. Some two factor interactions are confounded with four factor and higher order interactions.

In resolution VII designs, main effects, two-factor inter- actions, and three-factor interactions are mutually not confounded with one another.

Design resolution is intended to provide a quick means of assessing whether an experimental design allows for the estimation of all important effects (main effects and low-order interactions).

COMPLETELY RANDOMIZED FRACTIONAL FACTORIAL (2^{n-p}) DESIGNS

The number of test runs needed for complete factorial experiments increases rapidly as the number of factors increases, even if each factor is tested at only two levels. The intentional confounding of factor effects using fractional factorials can reduce the number of test runs substantially for large designs if some interaction effects are known to be nonexistent or negligible relative to the uncontrolled experimental error variation.

Designing a $\left(\frac{1}{2}\right)^p$ Fraction of a 2^n Design: 2^{n-p}

1. Select p defining contrasts

   None of these contrasts can be obtained by a multiplication of the other contrasts.

2. An additional $2^p - p - 1$ implicit defining contrasts are determined by multiplying the p contrasts selected in Step 1.

3. Randomly assign $+1$ or -1 to the p defining contrasts yielding a vector of all $+1$'s or all -1's

4. Select the Treatments having the specified p-vector for their values in the defining contrasts.

   Note, that there are 2^{n-p} Treatments selected for the experiment. Thus, the experimental data will yield estimates of $2^{n-p} - 1$ effects.

5. The Alias sets of confounded contrasts can be obtained by using the 2^p defining and implicit contrasts.

6. The n factors yield a total of 2^n effects:

   overall mean, n main effects, $\binom{n}{2}$ two-way interactions, $\binom{n}{3}$ three-way interactions, ...,

   one n-way interaction for a total of $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ effects

   The 2^n effects are categorized into 2^{n-p} alias sets, each containing 2^p elements.

   The effects in a given alias set are confounded.

The above ideas will be illustrated by considering specific fractional factorial designs:

Half Fractions: 2^{n-1} Designs : $p = 1$

Half fractions of factorial experiments are specified by first stating the defining contrast for the fraction. The defining contrast is the effect that is confounded with the constant effect. It can be expressed symbolically as an equation, the defining equation, by setting the confounded effect equal to I, which represents the constant effect. For the manufacturing-plant example, main half fractions of the complete 2^5 experiment can be identified in Table 2. We could combine any pair of the four blocks. For example, Blocks 1 and 2 would have a defining equations for the fractional factorial experiment $I = A_2A_3A_4A_5$. This would yield a design having main effects confounded with three-factor interactions.

– e.g. choose 2 of the 4 blocks so that we observe 16 of the 32 possible treatments

Ordinarily the effect chosen for the defining contrast is the highest-order interaction among the factors. Once the defining contrast is chosen, the half fraction of the factor-level combinations that are to be included in the experiment are those that have either the positive or the negative signs in the effects representation of the defining contrast. The choice of the positive or the negative signs is usually made randomly (e.g., by flipping a coin). As an illustration of this procedure, consider using $I = A_1A_2A_3A_4A_5$ in Table 2 on page 6. The treatments in Table 2 having +1 associated with $A_1A_2A_3A_4A_5$ are selected for use in the experiment.

Using $I = A_1A_2A_3A_4A_5 = +1$ in Table 2 yields the half-fraction given in Table 3.

Table 3: A 1/2 fraction of a 2^5 factorial experiment (2^{5-1}) Using $A_1A_2A_3A_4A_5$ as the Generator

TRT	DATA	A_1	A_2	A_3	A_4	A_5	$A_1A_2A_3A_4A_5$	$A_2A_3A_4A_5$
2	Y_{11112}	-1	-1	-1	-1	1	1	-1
3	Y_{11121}	-1	-1	-1	1	-1	1	-1
5	Y_{11211}	-1	-1	1	-1	-1	1	-1
8	Y_{11222}	-1	-1	1	1	1	1	-1
9	Y_{12111}	-1	1	-1	-1	-1	1	-1
12	Y_{12122}	-1	1	-1	1	1	1	-1
14	Y_{12212}	-1	1	1	-1	1	1	-1
15	Y_{12221}	-1	1	1	1	-1	1	-1
17	Y_{21111}	1	-1	-1	-1	-1	1	1
20	Y_{21122}	1	-1	-1	1	1	1	1
22	Y_{21212}	1	-1	1	-1	1	1	1
23	Y_{21221}	1	-1	1	1	-1	1	1
26	Y_{22112}	1	1	-1	-1	1	1	1
27	Y_{22121}	1	1	-1	1	-1	1	1
29	Y_{22211}	1	1	1	-1	-1	1	1
32	Y_{22222}	1	1	1	1	1	1	1

notice: $A_1 \cdot A_1A_2A_3A_4A_5 = I \cdot A_2A_3A_4A_5$

ESTIMATING MAIN AND INTERACTION EFFECTS IN 2^{n-p} DESIGNS

Suppose you have a 2^{n-p} design with n the number of factors: $A_1, A_2, \cdots, A_n$, $N = 2^{n-p}$ the number of data values, $Y_{ijk....}$'s

The following will describe how to estimate the effects in such a design:

1. **Main Effects:** Factor A_1 has only two levels hence the estimated main effect of Factor A_1 is the mean difference between the data values associated with the high level of A_1 and the data values associated with the low level of A_1, that is, mean of N/2 data values associated with +1 and the mean of N/2 data values associated with -1:

   Main Effect of Factor A_1: $\bar{Y}_{2....} - \bar{Y}_{1....}$ where

   $$\bar{Y}_{2....} - \bar{Y}_{1....} = \frac{1}{N/2}(\text{sum of N/2 Y's having} A_1 = +1) - \frac{1}{N/2}(\text{sum of N/2 Y's having} A_1 = -1)$$

   Example Suppose we have a 2^{5-1} design, $n = 5$ factors and $N = 2^{5-1} = 16$ data values with generator $I = A_1A_2A_3A_4A_5$

   Main Effect of Factor A_1: $\bar{Y}_{2....} - \bar{Y}_{1....}$ is given as follows using the coefficients in Table 3 on page 11 in Handout 10:

   $$\begin{aligned}\bar{Y}_{2....} - \bar{Y}_{1....} &= \frac{1}{8}(Y_{21111} + Y_{21122} + Y_{21212} + Y_{21221} + Y_{22112} + Y_{22121} + Y_{22211} + Y_{22222}) \\ &\quad -\frac{1}{8}(Y_{11112} + Y_{11121} + Y_{11211} + Y_{11222} + Y_{12111} + Y_{12122} + Y_{12212} + Y_{12221}) \\ &= \frac{2}{16}[(\text{sum of data with +1}) - (\text{sum of data with -1})]\end{aligned}$$

2. **Two-way Interactions:** The two-way interaction between Factors A_1 and A_2 is given by $(\bar{Y}_{11...} - \bar{Y}_{12...}) - (\bar{Y}_{21...} - \bar{Y}_{22...})$ where

   $$\begin{aligned}(\bar{Y}_{11...} - \bar{Y}_{12...}) - (\bar{Y}_{21...} - \bar{Y}_{22...}) &= \left(\frac{1}{N/4}(\text{sum of N/4 Y's having} A_1 = -1, A_2 = -1)\right) \\ &\quad - \left(\frac{1}{N/4}(\text{sum of N/4 Y's having} A_1 = -1, A_2 = +1)\right) \\ &\quad - \left(\frac{1}{N/4}(\text{sum of N/4 Y's having} A_1 = +1, A_2 = -1)\right) \\ &\quad + \left(\frac{1}{N/4}(\text{sum of N/4 Y's having} A_1 = +1, A_2 = +1)\right)\end{aligned}$$

   In our example, the two way interaction between A_1 and A_2 is given by

Two-way Interactions: A_1A_2:

$$
\begin{aligned}
(\bar{Y}_{11\ldots} - \bar{Y}_{12\ldots}) - (\bar{Y}_{21\ldots} - \bar{Y}_{22\ldots}) \quad = \quad & \frac{1}{4}\left(Y_{11112} + Y_{11121} + Y_{11211} + Y_{11222}\right) \\
& -\frac{1}{4}\left(Y_{12111} + Y_{12122} + Y_{12212} + Y_{12221}\right) \\
& -\frac{1}{4}\left(Y_{21111} + Y_{21122} + Y_{21212} + Y_{21221}\right) \\
& +\frac{1}{4}\left(Y_{22112} + Y_{22121} + Y_{22211} + Y_{22222}\right)
\end{aligned}
$$

3. **3-way and larger interactions** The k-way interactions are obtained in a similar fashion. The formula would just be

   $\frac{1}{N/2^k}$ times the differences in the various sums of responses.

DESIGNING HALF FRACTIONS OF TWO-LEVEL FACTORIAL EXPERIMENTS IN COMPLETELY RANDOMIZED DESIGNS

1. Choose the defining contrast: the effect that is to be confounded with the constant effect.
2. Randomly decide whether the experiment will contain the factor-level combinations with the positive or the negative signs in the effects representation of the defining contrast.
3. Form a table containing the effects representations of the main effects for each of the factors. Add a column containing the effects representation of the defining contrast.
4. Select the factor-level combinations that have the chosen sign in the effects representation of the defining contrast.
5. Randomize the assignment of the factor-level combinations to the experimental units or to the test sequence, as appropriate.

The resolution of a half fraction of a complete factorial experiment equals the number of factors included in the defining contrast. Half fractions of highest resolution, therefore, are those that confound the highest-order interaction with the constant effect. The highest resolution a half fraction can attain is equal to the number of factors in the experiment. The confounding pattern of a half fraction of a complete factorial is determined by symbolically multiplying each side of the defining equation by each of the factor effects. The procedure given below is used to determine the Confounding Pattern of effects.

$I = A_1A_2A_3A_4A_5$

$A_1 = A_1A_1A_2A_3A_4A_5 = A_2A_3A_4A_5$ | $A_1A_2 = A_3A_4A_5$

$A_2 = A_1A_3A_4A_5$

$\vdots$

1. Write the defining equation of the fractional factorial experiment.
2. Symbolically multiply both sides of the defining equation by one of the factor effects.
3. Reduce the right side of the equation using the following algebraic convention: For any factor, say X, $X * I = X$ and $X * X = X^2 = I$.
4. Repeat steps 2 and 3 until each factor effect is listed in either the left or the right side of one of the equations.

The symbolic multiplication of two effects is equivalent to multiplying the individual elements in the effects representations of the two effects. Note that the constant effect has all its elements equal to 1. Consequently, any effect multiplying the constant effect remains unchanged. This is the reason for the convention $X * I = X$. Similarly, any effect multiplying itself will result in a column of ones because each element is the square of either + 1 or -1. Consequently, an effect multiplying itself is a column of + 1 values: $X * X = I$. As an illustration of this procedure consider the half fraction displayed in Table 3. The defining equation is $I = A_1A_2A_3A_4A_5$. Multiplying this equation by each of the other effects, using the algebraic convention defined in step 4, results in the following confounding pattern for the factor effects:

Table 4: Alias Sets for the design in Table 3

Alias Set			Alias Set		
1	I	$A_1A_2A_3A_4A_5$	9	A_1A_4	$A_2A_3A_5$
2	A_1	$A_2A_3A_4A_5$	10	A_1A_5	$A_2A_3A_4$
3	A_2	$A_1A_3A_4A_5$	11	A_2A_3	$A_1A_4A_5$
4	A_3	$A_1A_2A_4A_5$	12	A_2A_4	$A_1A_3A_5$
5	A_4	$A_1A_2A_3A_5$	13	A_2A_5	$A_1A_3A_4$
6	A_5	$A_1A_2A_3A_4$	14	A_3A_4	$A_1A_2A_5$
7	A_1A_2	$A_3A_4A_5$	15	A_3A_5	$A_1A_2A_4$
8	A_1A_3	$A_2A_4A_5$	16	A_4A_5	$A_1A_2A_3$

See slide 10. We have 2^{n-p} alias sets each containing 2^p treatments.

By examining the above alias sets, we can observe that this is a resolution V design.

R=V because

1. Main effects (s=1) are not confounded with effects containing fewer than

   R-1=5-1=4 factors:

- main effects are not confounded with two factor interactions and three factor interactions but some of the main effects are confounded with four factor interactions.

2. Two factor interactions (s=2) are not confounded with effects containing fewer than

   R-2 = 5-2 = 3 factors:

- two factor interactions are not confounded with any other two factor interactions but some of the two factor interactions are confounded with three factor interactions.

3. Three factor interactions (s=3) are not confounded with effects containing fewer than

   R-2 = 5-3 = 2 factors:

- three factor interactions are not confounded with main effects. Some three factor interactions are confounded with two factor interactions.

Quarter (2^{n-2}) and Smaller Fractions (2^{n-p}, where $p > 2$)

Quarter and smaller fractions of two-level complete factorial experiments are constructed similarly to half fractions. The major distinction is that more than one defining contrast is needed to partition the factor-level combinations. Consider designing an experiment that is to include a quarter fraction of a five-factor complete factorial experiment. For example, suppose only 8 of the runs can be conducted for the acid-plant corrosion-rate study. By using two defining contrasts, $\frac{1}{4}$ of the factor-level combinations can be selected. Suppose the two defining contrasts chosen are $A_1A_2A_3$ and $A_1A_2A_3A_4A_5$ Suppose further that the negative sign is randomly assigned to both of the defining contrasts. By referring to Table 1 on page 3, we can observe that the factor-level combinations included in the experiment are those given in the Table 5.

$p=2$

Table 5: A 1/4 fraction of a 2^5 factorial experiment: 2^{5-2} Using $A_1A_2A_3$ and $A_1A_2A_3A_4A_5$ as the Generators

confounded.

TRT	A_1	A_2	A_3	A_4	A_5	$A_1A_2A_3$	$A_1A_2A_3A_4A_5$	A_4A_5
1	-1	-1	-1	-1	-1	-1	-1	1
4	-1	-1	-1	1	1	-1	-1	1
13	-1	1	1	-1	-1	-1	-1	1
16	-1	1	1	1	1	-1	-1	1
21	1	-1	1	-1	-1	-1	-1	1
24	1	-1	1	1	1	-1	-1	1
25	1	1	-1	-1	-1	-1	-1	1
28	1	1	-1	1	1	-1	-1	1

Although only two defining equations need to be chosen to select the factor-level combinations for a quarter fraction of a complete factorial experiment, a third equation is also satisfied. Note in the last column of Table 5 that all the factor-level combinations that satisfy the first two defining equations, $I_1 = -A_1A_2A_3$ and $I_2 = -A_1A_2A_3A_4A_5$, also satisfy $I_3 = A_4A_5$. A third defining equation is always satisfied when two defining contrasts are chosen. This third implicit defining contrast, I_3 can be identified by symbolically, multiplying the other two contrasts. In this example,

$$I_1 * I_2 = (-A_1A_2A_3) * (-A_1A_2A_3A_4A_5) = A_4A_5 = I_3$$

A concise way of expressing the defining contrasts is $I = -A_1A_2A_3 = -A_1A_2A_3A_4A_5 (= A_4A_5)$, with the implied contrast in parentheses being optional, since it can be easily be determined from the other two.

The resolution of quarter fractions of complete factorials constructed in this fashion equals the number of factors in the smallest of the defining contrasts, including the one implied by the two chosen for the design. Is the above a *good* design?

No, main effects of A_4 & A_5 are confounded.

Thus in the above example, the resulting quarter fraction is a resolution-II design because the smallest number of factors in the defining contrasts is two ($I = A_4A_5$). Just as there are three defining contrast in a quarter fraction of a complete factorial experiment, each of the factor effects in the experiment is confounded with three additional effects. The confounding pattern is determined by symbolically multiplying the defining equations, including the implied one, by each of the factor effects.

Suppose the defining equations for a quarter fraction of a five-factor factorial are

$$A_1A_2A_3 = A_1A_2A_3A_4A_5$$

Then, a concise way of expressing the defining contrasts is

$$I = A_1A_2A_3 = A_1A_2A_3A_4A_5 = A_4A_5$$

then the main effects are confounded as follows:

Multiple each of the Factors A_1, A_2, A_3, A_4, A_5 times the four terms in the above expression, yielding:

$$A_1 = A_2A_3 = A_2A_3A_4A_5 = A_1A_4A_5$$
$$A_2 = A_1A_3 = A_1A_3A_4A_5 = A_2A_4A_5$$
$$A_3 = A_1A_2 = A_1A_2A_4A_5 = A_3A_4A_5$$
$$A_4 = A_1A_2A_3A_4 = A_1A_2A_3A_5 = A_5$$
$$A_5 = A_1A_2A_3A_5 = A_1A_2A_3A_4 = A_4$$

This is not a very useful design since the main effects of factors A_4 and A_5 are in the same alias set and hence are confounded. The remaining alias sets are given in Table 6.

Table 6: Alias Sets for the design in Table 5

Alias Set				
1	I	$A_1A_2A_3$	$A_1A_2A_3A_4A_5$	A_4A_5
2	A_1	A_2A_3	$A_2A_3A_4A_5$	$A_1A_4A_5$
3	A_2	A_1A_3	$A_1A_3A_4A_5$	$A_2A_4A_5$
4	A_3	A_1A_2	$A_1A_2A_4A_5$	$A_3A_4A_5$
5	A_4	$A_1A_2A_3A_4$	$A_1A_2A_3A_5$	A_5
6	A_1A_4	$A_2A_3A_4$	$A_2A_3A_5$	A_1A_5
7	A_2A_4	$A_1A_3A_4$	$A_1A_3A_5$	A_2A_5
8	A_3A_4	$A_1A_2A_4$	$A_1A_2A_5$	A_3A_5

There are $2^{n-p} = 2^{5-2} = 2^3 = 8$ Alias sets w/ $2^p = 2^2 = 4$ TCTs in each set.

A higher-resolution fractional factorial design can be constructed by selecting the defining equation:

$$I = A_1A_2A_4 = A_1A_3A_5 \;\; (= A_2A_3A_4A_5)$$

This design is resolution III and is displayed in Table 7 with the alias sets given in Table 8.

Table 7: A 1/4 fraction of a 2^5 factorial experiment: 2^{5-2} Using $A_1A_2A_4$ and $A_1A_3A_5$

TRT	A_1	A_2	A_3	A_4	A_5	$A_1A_2A_4$	$A_1A_3A_5$	$A_2A_3A_4A_5$
4	-1	-1	-1	1	1	1	1	1
7	-1	-1	1	1	-1	1	1	1
10	-1	1	-1	-1	1	1	1	1
13	-1	1	1	-1	-1	1	1	1
17	1	-1	-1	-1	-1	1	1	1
22	1	-1	1	-1	1	1	1	1
27	1	1	-1	1	-1	1	1	1
32	1	1	1	1	1	1	1	1

Table 8: Alias Sets for design in Table 7

Alias Set				
1	I	$A_1A_2A_4$	$A_1A_3A_5$	$A_2A_3A_4A_5$
2	A_1	A_2A_4	A_3A_5	$A_1A_2A_3A_4A_5$
3	A_2	A_1A_4	$A_1A_2A_3A_5$	$A_3A_4A_5$
4	A_3	$A_1A_2A_3A_4$	A_1A_5	$A_2A_4A_5$
5	A_4	A_1A_2	$A_1A_3A_4A_5$	$A_2A_3A_5$
6	A_5	$A_1A_2A_4A_5$	A_1A_3	$A_2A_3A_4$
7	A_2A_3	$A_1A_3A_4$	$A_1A_2A_5$	A_4A_5
8	A_2A_5	$A_1A_4A_5$	$A_1A_2A_3$	A_3A_4

The general procedure for constructing fractional factorial experiments for two-level factors is consists of the following steps.

1. A 2^p fraction of a 2^n (a 2^{n-p} fractional factorial) requires that p defining contrasts be chosen, none of which is obtainable by algebraic multiplication of the other contrasts.

2. An additional $2^p - p - 1$ implicit defining contrasts are then determined by algebraically multiplying the p chosen contrasts.

3. The resolution of the resulting completely randomized fractional factorial experiment equals the number of factors in the smallest defining contrast, including the implicit ones.

4. The confounding pattern is obtained by multiplying the defining equations by the factor effects.

5. There are 2^{n-p} Alias Sets consisting of 2^p contrasts each.

6. The contrasts within a given Alias Set are confounded effects.

Table 12A.1 on page 421 of the textbook provides a list of fractional factorial experiments involving from 3 to 10 factors. Included in the table are the defining equations and the resolutions of the designs. Resolution-IV and resolution- V designs are sought because the primary interest in multifactor experiments is generally in the analysis of main effects and two-factor interactions.

Appendix: Fractional Factorial Design Plans

Statistical Principles of Research Design & Analysis, Author: Kuehl

* If fraction = 1/2 and # factors = K. always use $I = A_1 A_2 \ldots A_K$ as design generator.

Table 12A.1 Selected 2^{n-p} fractional factorial designs

Number of Factors	*Experimental Units*	*Fraction*	*Design Resolution*	*Design Generator**
$2^3 = 8$ 3	4	$\frac{1}{2}$	III	$C = AB$ ⇒ $I = ABC$
$2^4 = 16$ 4	8	$\frac{1}{2}$	IV	$D = ABC$ ⇒ $I = ABCD$
$2^5 = 32$ 5	8	$\frac{1}{4}$	III	$D = AB$ $E = AC$ ⇒ $I = ABD = ACE$
	16	$\frac{1}{2}$	V	$E = ABCD$ ⇒ $I = ABCDE$
$2^6 = 64$ 6	16	$\frac{1}{4}$	IV	$E = ABC$ $F = BCD$ ⇒ $I = ABC = BCD$
	8	$\frac{1}{8}$	III	$D = AB$ $E = AC$ $F = BC$
7	32	$\frac{1}{4}$	IV	$F = ABCD$ $G = ABDE$
	16	$\frac{1}{8}$	IV	$E = ABC$ $F = BCD$ $G = ACD$
	8	$\frac{1}{16}$	III	$D = AB$ $E = AC$ $F = BC$ $G = ABC$
8	64	$\frac{1}{4}$	V	$G = ABCD$ $H = ABEF$
	32	$\frac{1}{8}$	IV	$F = ABC$ $G = ABD$ $H = BCDE$
	16	$\frac{1}{16}$	IV	$E = BCD$ $F = ACD$ $G = ABC$ $H = ABD$

Table 12A.1 (*continued*)

Number of Factors	Experimental Units	Fraction	Design Resolution	Design Generator*
9	16	$\frac{1}{32}$	III	$E = ABC$
				$F = BCD$
				$G = ACD$
				$H = ABD$
				$J = ABCD$
	64	$\frac{1}{8}$	IV	$G = ABCD$
				$H = ACEF$
				$J = CDEF$
	32	$\frac{1}{16}$	IV	$F = BCDE$
				$G = ACDE$
				$H = ABDE$
				$J = ABCE$
10	16	$\frac{1}{64}$	III	$E = ABC$
				$F = BCD$
				$G = ACD$
				$H = ABD$
				$J = ABCD$
				$K = AB$
	32	$\frac{1}{32}$	IV	$F = ABCD$
				$G = ABCE$
				$H = ABDE$
				$J = ACDE$
				$K = BCDE$
	64	$\frac{1}{16}$	V	$G = BCDF$
				$H = AGDB$
				$J = ABDE$
				$K = ABCE$
	128	$\frac{1}{8}$	V	$H = ABCG$
				$J = BCDE$
				$K = ACDF$

* Either the positive or negative half of the design generators may be used to construct the fractional design.

SAS Code for Designing a Fractional Factorial Design

```
*fractional,main.sas ;

proc factex;
      factors A B C D E F G H/nlev=2;
      model resolution=maximum;
      size design=16;
      examine design confounding aliasing(4);
   run;

   OUTPUT FROM SAS:
```

$n=8 \Rightarrow 2^8$ design $= 256$
w/ 2 levels for each factor.

$p=4 \Rightarrow 2^8 \cdot 2^{-4} = 2^4 = 16$

Limiting interactions to at most 4-way.

The FACTEX Procedure

Factor Confounding Rules - Defining Contrasts (Generators)

```
E = B*C*D
F = A*C*D
G = A*B*D
H = A*B*C
```

$I = BCDE$
$I = ACDF$
$I = ABDG$
$I = ABCH$

see table on pg 21

Design Points

Experiment Number	A	B	C	D	E	F	G	H
1	-1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	1	1	1	1	-1
3	-1	-1	1	-1	1	1	-1	1
4	-1	-1	1	1	-1	-1	1	1
5	-1	1	-1	-1	1	-1	1	1
6	-1	1	-1	1	-1	1	-1	1
7	-1	1	1	-1	-1	1	1	-1
8	-1	1	1	1	1	-1	-1	-1
9	1	-1	-1	-1	-1	1	1	1
10	1	-1	-1	1	1	-1	-1	1
11	1	-1	1	-1	1	-1	1	-1
12	1	-1	1	1	-1	1	-1	-1
13	1	1	-1	-1	1	1	-1	-1
14	1	1	-1	1	-1	-1	1	-1
15	1	1	1	-1	-1	-1	-1	1
16	1	1	1	1	1	1	1	1

Let D_1 = BCDE, D_2 = ACDF, D_3 = ABDG, D_4 = ABCH

Q: where do these come from.

Aliasing Structure for up to 4-way Interactions

D_4 D_3 $-D_1 \cdot D_2$ D_2 $-D_1 \cdot D_3$ $-D_1 \cdot D_4$ $-D_1$ $-D_2 \cdot D_3$

0 = A*B*C*H = A*B*D*G = A*B*E*F = A*C*D*F = A*C*E*G = A*D*E*H = A*F*G*H = B*C*D*E = B*C*F*G

$D_2 \cdot D_4$ = B*D*F*H = B*E*G*H = C*D*G*H = C*E*F*H = D*E*F*G $D_3 \cdot D_4$

A = B*C*H = B*D*G = B*E*F = C*D*F = C*E*G = D*E*H = F*G*H

B = A*C*H = A*D*G = A*E*F = C*D*E = C*F*G = D*F*H = E*G*H

C = A*B*H = A*D*F = A*E*G = B*D*E = B*F*G = D*G*H = E*F*H

D = A*B*G = A*C*F = A*E*H = B*C*E = B*F*H = C*G*H = E*F*G

E = A*B*F = A*C*G = A*D*H = B*C*D = B*G*H = C*F*H = D*F*G

F = A*B*E = A*C*D = A*G*H = B*C*G = B*D*H = C*E*H = D*E*G

G = A*B*D = A*C*E = A*F*H = B*C*F = B*E*H = C*D*H = D*E*F

H = A*B*C = A*D*E = A*F*G = B*D*F = B*E*G = C*D*G = C*E*F

Q: she also said we get these by multiplying the letters on the LHS by each of the things on the top row, but that doesn't add up

A*B = C*H = D*G = E*F = A*C*D*E = A*C*F*G = A*D*F*H = A*E*G*H = B*C*D*F = B*C*E*G = B*D*E*H = B*F*G*H

A*C = B*H = D*F = E*G = A*B*D*E = A*B*F*G = A*D*G*H = A*E*F*H = B*C*D*G = B*C*E*F = C*D*E*H = C*F*G*H

A*D = B*G = C*F = E*H = A*B*C*E = A*B*F*H = A*C*G*H = A*E*F*G = B*C*D*H = B*D*E*F = C*D*E*G = D*F*G*H

A*E = B*F = C*G = D*H = A*B*C*D = A*B*G*H = A*C*F*H = A*D*F*G = B*C*E*H = B*D*E*G = C*D*E*F = E*F*G*H

A*F = B*E = C*D = G*H = A*B*C*G = A*B*D*H = A*C*E*H = A*D*E*G = B*C*F*H = B*D*F*G = C*E*F*G = D*E*F*H

A*G = B*D = C*E = F*H = A*B*C*F = A*B*E*H = A*C*D*H = A*D*E*F = B*C*G*H = B*E*F*G = C*D*F*G = D*E*G*H

A*H = B*C = D*E = F*G = A*B*D*F = A*B*E*G = A*C*D*G = A*C*E*F = B*D*G*H = B*E*F*H = C*D*F*H = C*E*G*H

GENERATING THE ALIAS SETS IN R

The following R code, **FractionalFact.R** will produce the design matrix and alias sets that were produced by SAS for a 8 factor design in 16 runs using the generators

I = BCDE = ACDF = ABDG = ABCH or equivalently, E=BCD, F=ACD, G=ABD, H=ABC

This will yield a resolution IV design, R = IV:

```
install.packages("FrF2")
library(FrF2)
design = FrF2(nruns=16,nfactors=8,generators=c("BCD","ACD","ABD","ABC"))
design$y = c(1:nrow(design))
alias_sets = aliases(lm(y~(.)^4,data=design))
class=design
alias_sets

#R output
```

```
 Alias Sets:

 A = B:C:H = B:D:G = B:E:F = C:D:F = C:E:G = D:E:H = F:G:H
 B = A:C:H = A:D:G = A:E:F = C:D:E = C:F:G = D:F:H = E:G:H
 C = A:B:H = A:D:F = A:E:G = B:D:E = B:F:G = D:G:H = E:F:H
 D = A:B:G = A:C:F = A:E:H = B:C:E = B:F:H = C:G:H = E:F:G
 E = A:B:F = A:C:G = A:D:H = B:C:D = B:G:H = C:F:H = D:F:G
 F = A:B:E = A:C:D = A:G:H = B:C:G = B:D:H = C:E:H = D:E:G
 G = A:B:D = A:C:E = A:F:H = B:C:F = B:E:H = C:D:H = D:E:F
 H = A:B:C = A:D:E = A:F:G = B:D:F = B:E:G = C:D:G = C:E:F
 A:B = C:H = D:G = E:F = A:C:D:E = A:C:F:G = A:D:F:H = A:E:G:H = B:C:D:F = B:C:E:G = B:D:E:H = B:F:G:H
 A:C = B:H = D:F = E:G = A:B:D:E = A:B:F:G = A:D:G:H = A:E:F:H = B:C:D:G = B:C:E:F = C:D:E:H = C:F:G:H
 A:D = B:G = C:F = E:H = A:B:C:E = A:B:F:H = A:C:G:H = A:E:F:G = B:C:D:H = B:D:E:F = C:D:E:G = D:F:G:H
 A:E = B:F = C:G = D:H = A:B:C:D = A:B:G:H = A:C:F:H = A:D:F:G = B:C:E:H = B:D:E:G = C:D:E:F = E:F:G:H
 A:F = B:E = C:D = G:H = A:B:C:G = A:B:D:H = A:C:E:H = A:D:E:G = B:C:F:H = B:D:F:G = C:E:F:G = D:E:F:H
 A:G = B:D = C:E = F:H = A:B:C:F = A:B:E:H = A:C:D:H = A:D:E:F = B:C:G:H = B:E:F:G = C:D:F:G = D:E:G:H
 A:H = B:C = D:E = F:G = A:B:D:F = A:B:E:G = A:C:D:G = A:C:E:F = B:D:G:H = B:E:F:H = C:D:F:H = C:E:G:H
```

class

```
#R output - Design Matrix - Which Treatments to use in the 16 Runs

Run  A  B  C  D  E  F  G  H  y   Treatment
 1   1  1  1  1  1  1  1  1  1   All 8 factors at the High Level
 2  -1 -1 -1  1  1  1  1 -1  2   A,B,C,H at Low Level and rest at High Level
 3  -1 -1 -1 -1 -1 -1 -1 -1  3   All 8 factors at the Low Level
 4  -1 -1  1  1 -1 -1  1  1  4   A,B,E,F at Low Level and rest at High Level
 5   1  1  1 -1 -1 -1 -1  1  5   D,E,F,G at Low Level and rest at High Level
 6  -1  1  1 -1 -1  1  1 -1  6   A,D,E,H at Low Level and rest at High Level
 7  -1  1  1  1  1 -1 -1 -1  7   A,F,G,H at Low Level and rest at High Level
 8   1 -1 -1  1  1 -1 -1  1  8   B,C,F,G at Low Level and rest at High Level
 9   1 -1  1  1 -1  1 -1 -1  9   B,E,G,H at Low Level and rest at High Level
10   1 -1 -1 -1 -1  1  1  1 10   B,C,D,E at Low Level and rest at High Level
11  -1  1 -1  1 -1  1 -1  1 11   A,C,E,G at Low Level and rest at High Level
12   1  1 -1 -1  1  1 -1 -1 12   C,D,G,H at Low Level and rest at High Level
13   1  1 -1  1 -1 -1  1 -1 13   C,E,F,H at Low Level and rest at High Level
14  -1 -1  1 -1  1  1 -1  1 14   A,B,D,G at Low Level and rest at High Level
15   1 -1  1 -1  1 -1  1 -1 15   B,D,F,H at Low Level and rest at High Level
16  -1  1 -1 -1  1 -1  1  1 16   A,C,D,F at Low Level and rest at High Level
```


STOP Wednesday 4/13/22 (Week ~~3~~ 12 Lecture 32)

START Monday 4/18/22 (Week 13, Lecture 33)

SCREENING DESIGNS AND SEQUENTIAL EXPERIMENTATION

Screening experiments are conducted when a large number of factors are to be investigated but limited resources mandate that only a few runs be conducted. Screening experiments are conducted in order to identify a small number of dominant factors, often with the intent to conduct a more extensive investigation involving only these dominant factors. An important application of screening experiments is with *ruggedness tests* for laboratory test methods. The purpose of a ruggedness test is to determine environmental factors or test conditions (technicians, equipment fluctuations, maintenance cycles, etc.) that influence measurements obtained from the test method. Once these factors are identified, either they are tightly controlled or the test method is changed to reduce or eliminate their effects. The resulting test method then retains sensitivity, accuracy, and precision in spite of normal changes in operating conditions. One major goal of ruggedness testing is to permit the test method to be reliably used in different laboratories. Ruggedness testing and the application of screening designs can be extended to product and manufacturing process design. A rugged product or process possesses the .ability to withstand a wide variety of uses and conditions. Fractional factorial experiments are most often used as screening designs. Resolution- III designs allow the estimation of main effects that are not confounded with one another. Often R_{III} designs result in experiments that clearly identify the dominant factors without the major time and expense that would be required for higher-resolution designs.

A special class of fractional factorial experiments that is widely used in screening experiments was proposed by Plackett and Burman. These experiments have resolution III when conducted in completely randomized designs and are often referred to as *Plackett-Burman designs*. The designs discussed by Plackett and Burman are available for experiments that have the number of test runs equal to a multiple of four. Table 9A.4 on the next page lists design generators for experiments having 12, 16, 20, 24, and 32 runs.

The rows of the table denote the design factors, and the elements in each column are coded factor levels: a minus sign denotes one level of a factor and a plus sign denotes the other factor level. To allow for an estimate of experimental error, it is recommended that the design have at least six more test runs than the number of factors included in the experiment. The particular design selected depends on both the number of factors in the experiment and the number of runs. Each design generator can be used to construct a screening design having up to one fewer factor than the number of test runs. For example, the 16-run design generator can be used to construct screening designs for 1-15 factors; however, we again recommend that no more than 10 factors be used with a 16-run design. An illustration of the construction of screening experiments using the design generators is given on the next page.

If, as recommended, fewer factors than test runs are included in the design, several columns of the generated design are discarded. Any of the columns can be eliminated. Elimination can be based on experimental constraints or can be randomly selected. All rows (runs) for the retained columns are included in the design. The acid-plant corrosion-rate study was actually conducted as a screening experiment. To minimize the possibility of bias effects due to the run order, the experimental test sequence should be randomize as should the assignment of the factors to the columns. Screening designs often lead to further experimentation once dominant factors are identified. More generally, one often can plan experiments so that preliminary results determine the course of the experimentation.

Sequential experimentation is highly advisable when one embarks on a new course of experimental work in which little is known about the effects of any of the factors on the response. It would be unwise, for example, to design a complete factorial experiment in seven or more factors in such circumstances. The result of such a large test program might be that a small number of main effects and two-factor interactions are dominant. If so, a great amount of time and effort would have been wasted when a much smaller effort could have achieved the same results. If an experiment is contemplated in which a large number of factors are to be investigated, a key question that should be asked is whether the project goals will be met if only the dominant factor effects are identified. If so, a screening experiment should be performed, followed by a complete or fractional factorial experiment involving only the dominant factors. If not a resolution-V fractional factorial experiment would suffice.

Plackett - Burman Design Table and Example

	Number of Test Runs					
Factor Number	12	16	20	24	32	Factor Number
1	+	+	+	+	-	1
2	+	-	+	+	-	2
3	-	-	-	+	-	3
4	+	-	-	+	-	4
5	+	+	+	+	+	5
6	+	-	+	-	-	6
7	-	-	+	+	+	7
8	-	+	+	-	-	8
9	-	+	-	+	+	9
10	+	-	+	+	+	10
11	-	+	-	-	+	11
12		-	+	-	-	12
13		+	-	+	+	13
14		+	-	+	+	14
15		+	-	-	-	15
16			-	-	-	16
17			+	+	-	17
18			+	-	+	18
19			-	+	+	19
20				-	+	20
21				-	+	21
22				-	+	22
23				-	-	23
24					-	24
25					+	25
26					+	26
27					-	27
28					+	28
29					-	29
30					-	30
31					+	31

We will illustrate the construction and analysis of a Plackett-Burman Design using the following example.

Certain types of plated materials (for example, a connector in an electronic assembly) are soldered for strength. The manufacturer of the plated material tests it for solderability. Ten factors which may affect solderability are listed below:

Factor		Levels	
Factor Number	Name	-1	+1
1	Solvent Dip	No	Yes
2	Surface Area	Small	Large
3	Dip Device	Manual	Mechanical
4	Magnification	10 x	30 x
5	Solder Age	Fresh	Used
6	Flux Time	8 Sec	30 Sec
7	Drain Time	10 Sec	60 Sec
8	Stir	No	Yes
9	Solder Time	2 Sec	8 Sec
10	Residual Flux	Not Clean	Clean

The following table will identify which treatments to use in each of the 16 runs:

	Factor														
Runs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	+1	-1	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1
2	-1	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1
3	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1
4	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1	-1
5	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1	-1	-1
6	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1	-1	-1	+1
7	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1	-1	-1	+1	-1
8	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1	-1	-1	+1	-1	-1
9	+1	-1	+1	-1	+1	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1
10	-1	+1	-1	+1	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	+1
11	+1	-1	+1	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	+1	-1
12	-1	+1	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	+1	-1	+1
13	+1	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1
14	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1
15	+1	+1	-1	-1	-1	+1	-1	-1	+1	+1	-1	+1	-1	+1	+1
16	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

The factors column are then randomized and the runs are also randomized. This yields the following table for our example in which we ignore columns 11 through 15 because we have only 10 factors.

Plackett - Burman Design Example Cont.

Responses and Coded Factor Levels for Solder-Coverage Ruggedness Testing Data

		Factor									
~~Treatment~~ Runs	Solder Coverage(%)	1	2	3	4	5	6	7	8	9	10
1	91	1	-1	-1	-1	1	-1	-1	1	1	-1
2	97	-1	-1	-1	1	-1	-1	1	1	-1	1
3	89	-1	-1	1	-1	-1	1	1	-1	1	-1
4	82	-1	1	-1	-1	1	1	-1	1	-1	1
5	82	1	-1	-1	1	1	-1	1	-1	1	1
6	74	-1	-1	1	1	-1	1	-1	1	1	1
7	54	-1	1	1	-1	1	-1	1	1	1	1
8	66	1	1	-1	1	-1	1	1	1	1	-1
9	79	1	-1	1	-1	1	1	1	1	-1	-1
10	25	-1	1	-1	1	1	1	1	-1	-1	-1
11	77	1	-1	1	1	1	1	-1	-1	-1	1
12	44	-1	1	1	1	1	-1	-1	-1	1	-1
13	86	1	1	1	1	-1	-1	-1	1	-1	-1
14	97	1	1	1	-1	-1	-1	1	-1	-1	1
15	84	1	1	-1	-1	-1	1	-1	-1	1	1
16	97	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

The above data is then analyzed using a main effects only model:

- Cannot include 2-way effects, would confound if we included 2-way effects.
- Cant run w/ all 16 factors b/c then $t=16$, $n=16$

  $df_E = n - t - 1 = -1$

```
 ods html; ods graphics on;
* placket_burman.sas
* Analysis of Plackett-Burman  fractional factorial design;
option ls=75 ps=55 nocenter nodate formdlim='*';
data chem;
input run $ f1 $ f2 $ f3 $ f4 $ f5 $ f6 $ f7 $ f8  $ f9 $ f10  $ Y;
cards;
      1  +1    -1   -1   -1   +1   -1   -1   +1   +1    -1  91
      2  -1    -1   -1   +1   -1   -1   +1   +1   -1    +1  97
      3  -1    -1   +1   -1   -1   +1   +1   -1   +1    -1  89
      4  -1    +1   -1   -1   +1   +1   -1   +1   -1    +1  82
      5  +1    -1   -1   +1   +1   -1   +1   -1   +1    +1  82
      6  -1    -1   +1   +1   -1   +1   -1   +1   +1    +1  74
      7  -1    +1   +1   -1   +1   -1   +1   +1   +1    +1  54
      8  +1    +1   -1   +1   -1   +1   +1   +1   +1    -1  66
      9  +1    -1   +1   -1   +1   +1   +1   +1   -1    -1  79
     10  -1    +1   -1   +1   +1   +1   +1   -1   -1    -1  25
     11  +1    -1   +1   +1   +1   +1   -1   -1   -1    +1  77
     12  -1    +1   +1   +1   +1   -1   -1   -1   +1    -1  44
     13  +1    +1   +1   +1   -1   -1   -1   +1   -1    -1  86
     14  +1    +1   +1   -1   -1   -1   +1   -1   -1    +1  97
     15  +1    +1   -1   -1   -1   +1   -1   -1   +1    +1  84
     16  -1    -1   -1   -1   -1   -1   -1   -1   -1    -1  97
run;
proc print;
run;
proc glm;
class f1-f10;
model Y = f1-f10;
estimate 'f1' f1 1 -1;
estimate 'f2' f2 1 -1;
estimate 'f3' f3 1 -1;
estimate 'f4' f4 1 -1;
estimate 'f5' f5 1 -1;
estimate 'f6' f6 1 -1;
estimate 'f7' f7 1 -1;
estimate 'f8' f8 1 -1;
estimate 'f9' f9 1 -1;
estimate 'f10' f10 1 -1;
run;
data contrast;
input effects @@;
cards;
12.5 -18.5 -3 -15.25 -19.5 -9 -5.75 4.25 -7 8.75
run;
proc print; var effects;
proc rank out=new1 normal=blom;
var effects; ranks nquant; label nquant = 'normal quantiles';
proc gplot; plot effects*nquant;
title 'Placket Burman Design'; title2 'Normal Probability Plot of Effects';
run;
ods graphics off;ods html close;
```

* To get the estimate for main effect of f_1, average over the Y's w/ (+) in the f_1 column and then average over the (−) in the f_1 column and take the difference of the two

Main f_1:

$$\left(\frac{91+82+66+79+77+86+97+84}{8}\right) - \left(\frac{97+89+82+74+54+25+44+97}{8}\right)$$

$= 12.5$

[* matches estimate on pg 32]

This then yields the following AOV Table:

$F = \frac{MS_{RW}}{MSE}$ since all fixed effects.

ANOVA TABLE FOR Solder-Coverage Ruggedness TestS

	Source	DF	Sum of Squares	Mean Squares	F-statistic	p-value	
F_1	Solvent Dip	1	625.00	625.00	4.73	0.0815	$= P(F_{1,5} > 4.73)$
F_2	Surface Area	1	1369.00	1369.00	10.37	0.0235	
	Dip Device	1	36.00	36.00	0.27	0.6238	
F_4	Magnification	1	930.25	930.25	7.05	0.0452	
F_5	Solder Age	1	1521.00	1521.00	11.52	0.0194	
	Flux Time	1	324.00	324.00	2.45	0.1780	
	Drain Time	1	132.25	132.25	1.00	0.3628	
	Stir	1	72.25	72.25	0.55	0.4927	
	Solder Time	1	196.00	196.00	1.48	0.2774	
	Residual Flux	1	306.25	306.25	2.32	0.1882	
	Error	5	660.00	132.00			
	Total	15	6172.00				

In analyzing screening designs, it is often a good idea to allow a larger than normal significance level in testing for main effects when the goal of the screening experiment is to

Select potentially important factors for study in future experiments or

Identify factors for which tighter controls will be implemented in order to obtain a more uniform product

Using $\alpha = .10$ in the above AOV table, the factors requiring great control in order to achieve more consistent solderability test results would be Solvent Dip, Surface Area, Magnification, Solder Age.

Blacket Burman Design

Class	Levels	Values
f1	2	+1 -1
f2	2	+1 -1
f3	2	+1 -1
f4	2	+1 -1
f5	2	+1 -1
f6	2	+1 -1
f7	2	+1 -1
f8	2	+1 -1
f9	2	+1 -1
f10	2	+1 -1

Number of Observations Read 16
Number of Observations Used 16
Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	5512.000000	551.200000	4.18	0.0641
Error	5	660.000000	132.000000		
Corrected Total	15	6172.000000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
f1	1	625.000000	625.000000	4.73	0.0815
f2	1	1369.000000	1369.000000	10.37	0.0235
f3	1	36.000000	36.000000	0.27	0.6238
f4	1	930.250000	930.250000	7.05	0.0452
f5	1	1521.000000	1521.000000	11.52	0.0194
f6	1	324.000000	324.000000	2.45	0.1780
f7	1	132.250000	132.250000	1.00	0.3628
f8	1	72.250000	72.250000	0.55	0.4927
f9	1	196.000000	196.000000	1.48	0.2774
f10	1	306.250000	306.250000	2.32	0.1882

Parameter	Estimate	Standard Error	t Value	Pr > |t|
f1	12.5000000	5.74456265	2.18	0.0815
f2	-18.5000000	5.74456265	-3.22	0.0235
f3	-3.0000000	5.74456265	-0.52	0.6238
f4	-15.2500000	5.74456265	-2.65	0.0452
f5	-19.5000000	5.74456265	-3.39	0.0194
f6	-9.0000000	5.74456265	-1.57	0.1780
f7	-5.7500000	5.74456265	-1.00	0.3628
f8	4.2500000	5.74456265	0.74	0.4927
f9	-7.0000000	5.74456265	-1.22	0.2774
f10	8.7500000	5.74456265	1.52	0.1882

Still need to include F_3 in model.

<u>note</u>: when running model w/ interaction F_1*F_3, say F_1 is significant, F_3 is not and F_1*F_3 is.

Placket Burman Design

Normal Probability Plot of Effects

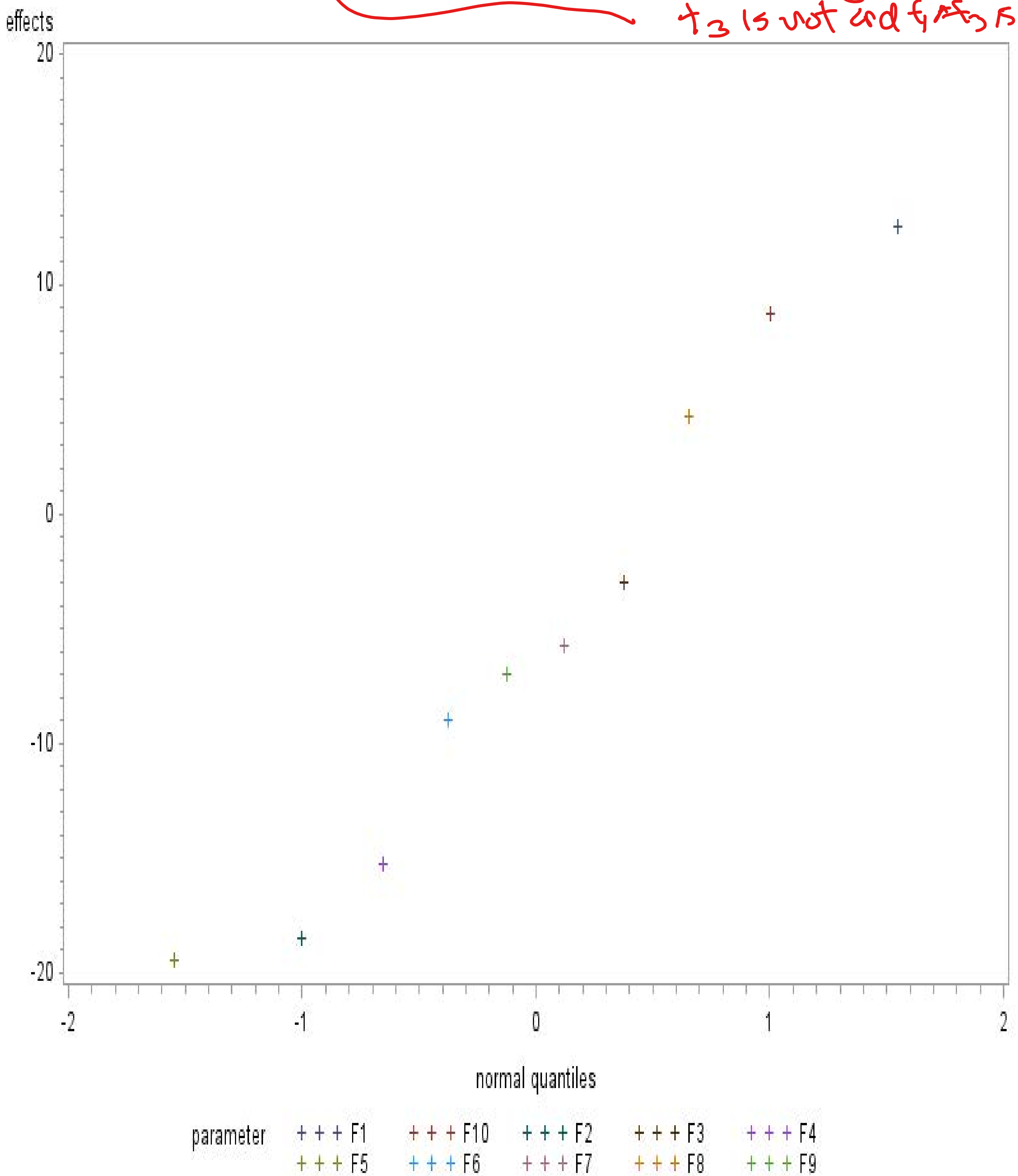


Illustration of the Analysis of a 2^{n-p} CRD Experiment

The analysis of a fractional factorial design will be illustrated using the following example from *Statistics for Experimentation*, by Box, Hunter, and Hunter.

In an injection molding experiment, eight variables were studied in a 2^{8-4} (a $\frac{1}{16}$ of a 2^8 factorial design). The response variable was the amount of shrinkage in the mold after the raw material was injected. The eight factors in Table 9 were identified as possibly affecting the amount of shrinkage.

Table 9: Injection Molding 2^{8-4} Fractional Factorial Experiment

Run	F_1 Mold Temp.	F_2 Moisture Content	F_3 Holding Pressure	F_4 Cavity Thickness	F_5 Booster Pressure	F_6 Cycle Time	F_7 Gate Size	F_8 Screw Speed	Shrinkage Y
1	-1	-1	-1	+1	+1	+1	-1	+1	14.0
2	+1	-1	-1	-1	-1	+1	+1	+1	16.8
3	-1	+1	-1	-1	+1	-1	+1	+1	15.0
4	+1	+1	-1	+1	-1	-1	-1	+1	15.4
5	-1	-1	+1	+1	-1	-1	+1	+1	27.6
6	+1	-1	+1	-1	+1	-1	-1	+1	24.0
7	-1	+1	+1	-1	-1	+1	-1	+1	27.4
8	+1	+1	+1	+1	+1	+1	+1	+1	22.6
9	+1	+1	+1	-1	-1	-1	+1	-1	22.3
10	-1	+1	+1	+1	+1	-1	-1	-1	17.1
11	+1	-1	+1	+1	-1	+1	-1	-1	21.5
12	-1	-1	+1	-1	+1	+1	+1	-1	17.5
13	+1	+1	-1	-1	+1	+1	-1	-1	15.9
14	-1	+1	-1	+1	-1	+1	+1	-1	21.9
15	+1	-1	-1	+1	+1	-1	+1	-1	16.7
16	-1	-1	-1	-1	-1	-1	-1	-1	20.3

Table 10 contains the calculated contrasts for the main effects and two-factor interactions.

Table 10: Estimated Contrasts

Effect	Estimator	Aliased Two-Factor Interactions
F_1	-0.7	
F_2	-0.1	
F_3	5.5	
F_4	-0.3	
F_5	-3.8	
F_6	-0.1	
F_7	0.6	
F_8	1.2	
F_1F_2	-1.2	$F_3 * F_7$, $F_4 * F_8$, $F_5 * F_6$
F_1F_3	1.8	$F_2 * F_7$, $F_4 * F_6$, $F_5 * F_8$
F_1F_4	-0.8	$F_2 * F_8$, $F_3 * F_6$, $F_5 * F_7$
F_1F_5	9.2	$F_2 * F_6$, $F_3 * F_8$, $F_4 * F_7$
F_1F_6	-0.6	$F_2 * F_5$, $F_3 * F_4$, $F_7 * F_8$
F_1F_7	-0.4	$F_2 * F_3$, $F_6 * F_8$, $F_4 * F_5$
F_1F_8	-1.2	$F_2 * F_4$, $F_3 * F_5$, $F_6 * F_7$

even though F_1 doesn't have a large effect when we are running this design we need to have F_3, F_5, F_1 & F_5 ; F_1 included b/c F_1 is a part of a significant interaction.

all the two way effects are confounded w/ other 2-way effects.

The SAS output provides a normal probability plot of the estimated effects. If the effects were all 0 then we would expect a straight-line fit to the 15 estimated responses since they would be simply a random sample from a normal distribution with mean 0 and variance σ_e. From the plot, we determine that estimated contrasts for the main effects of F_3 and F_5, and the two-factor interaction F_1F_5 appear to be different from 0. We cannot distinguish between the effect of the two-factor interaction F_1F_5 and the other three two-factor interactions: F_2F_6, F_3F_8, F_4F_7 since they are confounded.

The Main Effects can be calculated using the formulas from page 4 as follows:

Main Effect of F_1 : $\bar{Y}_+ - \bar{Y}_- \;=\; \mathbf{F}_1^{'}\mathbf{Y}/(N/2) \;=\; \mathbf{F}_1^{'}\mathbf{Y}/(8) \;=$

$[(-1)(14.0)+(1)(16.8)+(-1)(15.0)+(1)(15.4)+(-1)(27.6)+(1)(24.0)+(-1)(27.4)+(1)(22.6)+(1)(22.3)+$

$(-1)(17.1)+(1)(21.5)+(-1)(17.5)+(1)(15.9)+(-1)(21.9)+(1)(16.7)+(-1)(20.3)]/8 \;=\; -0.7$

The 2-way Interaction Effects can be calculated using the formulas from page 4 as follows:

2-way Interaction Effect of $F_1 * F_2$:

$(\bar{Y}_{++} - \bar{Y}_{+-}) - (\bar{Y}_{-+} - \bar{Y}_{--}) \;=\; (\mathbf{F}_1 \circ \mathbf{F}_2)^{'}\mathbf{Y}/(N/4) \;=\; (\mathbf{F}_1 \circ \mathbf{F}_2)^{'}\mathbf{Y}/(4) \;=$

$[(-1)(-1)(14.0)+(1)(-1)(16.8)+(-1)(1)(15.0)+(1)(1)(15.4)+(-1)(-1)(27.6)+(1)(-1)(24.0)+(-1)(1)(27.4)+$

$(1)(1)(22.6)+(1)(1)(22.3)+(-1)(1)(17.1)+(1)(-1)(21.5)+(-1)(-1)(17.5)+(1)(1)(15.9)+(-1)(1)(21.9)+$

$(1)(-1)(16.7)+(-1)(-1)(20.3)]/4 \;=\; -1.2$

SAS Code for Example:

```
*ffact_inj.sas;
option ls=75 ps=55 nocenter nodate;
data chem;
input v1  v2  v3  v4  v5  v6  v7  v8   Y;
cards;
L L L H H H L H 14.0
H L L L L H H H 16.8
L H L L H L H H 15.0
H H L H L L L H 15.4
L L H H L L H H 27.6
H L H L H L L H 24.0
L H H L L H L H 27.4
H H H H H H H H 22.6
H H H L L L H L 22.3
L H H H H L L L 17.1
H L H H L H L L 21.5
L L H L H H H L 17.5
H H L L H H L L 15.9
L H L H L H H L 21.9
H L L H H L H L 16.7
L L L L L L L L 20.3
run;
proc glm;
class v1-v8;
model Y = v1-v8 v1*v2 v1*v3 v1*v4 v1*v5 v1*v6 v1*v7 v1*v8;
estimate ‘v1’ v1 1 -1;
estimate ‘v2’ v2 1 -1;
estimate ‘v3’ v3 1 -1;
estimate ‘v4’ v4 1 -1;
estimate ‘v5’ v5 1 -1;
estimate ‘v6’ v6 1 -1;
estimate ‘v7’ v7 1 -1;
estimate ‘v8’ v8 1 -1;
estimate ‘v1*v2’ v1*v2 1 -1 -1  1;
estimate ‘v1*v3’ v1*v3 1 -1 -1  1;
estimate ‘v1*v4’ v1*v4 1 -1 -1  1;
estimate ‘v1*v5’ v1*v5 1 -1 -1  1;
estimate ‘v1*v6’ v1*v6 1 -1 -1  1;
estimate ‘v1*v7’ v1*v7 1 -1 -1  1;
estimate ‘v1*v8’ v1*v8 1 -1 -1  1;
run;
data contrast; input effects ;
cards;
-0.7  -0.1   5.5   -0.3   -3.8  -0.1   0.6  1.2
-1.2   1.8  -0.8    9.2   -0.6  -0.4  -1.2
run;
proc rank out=new1 normal=blom; var effects; ranks nquant;
label nquant = ’normal quantiles’;
proc gplot;
plot effects*nquant;
title ’Injection Molding’;  title2 ’Normal Probability Plot of Effects’;
run;
```

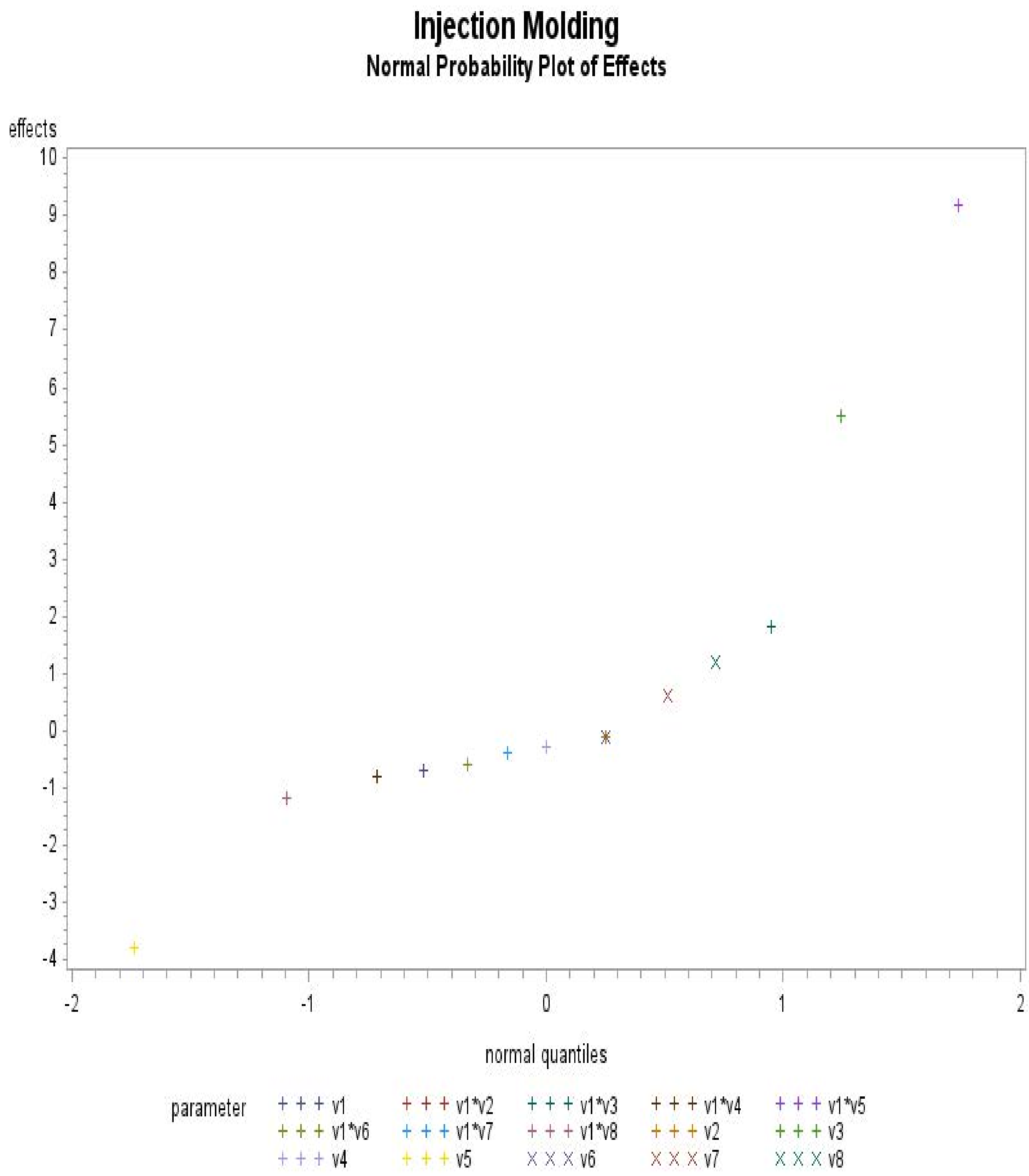
SAS OUTPUT:

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	280.280000	18.685333	.	.
Error	0	.	.		
Corrected Total	15	280.280000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
V1	1	1.960000	1.960000	.	.
V2	1	0.040000	0.040000	.	.
V3	1	121.000000	121.000000	.	.
V4	1	0.360000	0.360000	.	.
V5	1	57.760000	57.760000	.	.
V6	1	0.040000	0.040000	.	.
V7	1	1.440000	1.440000	.	.
v8	1	5.760000	5.760000	.	.
v1*v2	1	1.440000	1.440000	.	.
v1*v3	1	3.240000	3.240000	.	.
v1*v4	1	0.640000	0.640000	.	.
v1*v5	1	84.640000	84.640000	.	.
v1*v6	1	0.360000	0.360000	.	.
v1*v7	1	0.160000	0.160000	.	.
v1*v8	1	1.440000	1.440000	.	.

Parameter	Estimate	Standard Error	t Value	Pr > |t|
v1	-0.70000000	.	.	.
v2	-0.10000000	.	.	.
v3	5.50000000	.	.	.
v4	-0.30000000	.	.	.
v5	-3.80000000	.	.	.
v6	-0.10000000	.	.	.
v7	0.60000000	.	.	.
v8	1.20000000	.	.	.
v1*v2	-1.20000000	.	.	.
v1*v3	1.80000000	.	.	.
v1*v4	-0.80000000	.	.	.
v1*v5	9.20000000	.	.	.
v1*v6	-0.60000000	.	.	.
v1*v7	-0.40000000	.	.	.
v1*v8	-1.20000000	.	.	.



Injection Molding
Normal Probability Plot of Effects
effects
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3
-4
-2
-1
0
1
2
normal quantiles
parameter
v1
v1*v2
v1*v3
v1*v4
v1*v5
v1*v6
v1*v7
v1*v8
v2
v3
v4
v5
v6
v7
v8

3^n and Mixed Level Factorial Designs

Constructing fractional factorial designs when the desired number of levels is greater than two levels is considerably more involved than designs for two levels. Several good references for 3^{n-p} designs are

1. *Design and Analysis of Experiments*, by Douglas Montgomery

2. *Applied Factorial and Fractional Designs*, by V.L. Anderson and R.A. McLean

3. *Statistics for Experiments*, by G. Box, W. Hunter, J. Hunter

For designs having levels greater than 2 or 3, it is difficult to say much in general about their properties with respect to fractional factorials. It is necessary to evaluate each of the designs individually.

It is possible to extend the properties of 2^{n-p} designs to designs involving factors having number of levels equal to a power of 2. When examining such designs, it is important to carefully consider how to construct contrasts to represent main effects and interaction effects. The following example will illustrate some of the construction ideas involved in mixed levels factors where the levels are integer multiples of 2.

Suppose the experimenter is interested in investigating three factors:

A at four levels, B and C at two levels.

One approach is to define two new factors A_1 and A_2 each having two levels to represent the four levels of A:

$$(A_1, A_2) = \begin{cases} (-,-) & if \quad A = 1 \\ (-,+) & if \quad A = 2 \\ (+,-) & if \quad A = 3 \\ (+,+) & if \quad A = 4 \end{cases}$$

The complete factorial experiment would have $t = (4)(2)(2) = 16$ treatments. A half-fraction of the complete treatment combination could be constructed by using the defining equation:

$$I = A_1 A_2 BC$$

Using the methodology for 2^{n-p} designs, we would be inclined to think that the design is of Resolution IV, because four factors are involved in the defining equation. However, A_1A_2 represents one of the main effects for the four-level factor A and not two unique factors. Thus, the design is not resolution IV but resolution III. This can be more clearly seen by examining the confounding pattern for the main effects:

Alias Sets for the three factor design

Alias Set		
1	I	A_1A_2BC
2	A_1	A_2BC
3	A_2	A_1BC
4	A_1A_2	BC
5	B	A_1A_2C
6	C	A_1A_2B
7	A_1B	A_2C
8	A_2B	A_1C

Recall that (A_1, A_2) represents factor A having 4 levels. Therefore, there are three main effects for factor A which are represented by

$$A_1, \quad A_2, \quad A_1A_2$$

Thus, one of the three main effects for factor A, represented by A_1A_2 is confounded with the two-way interaction BC. Also, the other main effects A_1, A_2, B, and C are confounded with two-way or three-way interactions. Therefore, the design is of Resolution III.

General Method of Obtaining Aliasing

We obtained the alias structure (confounding groups) for 2^{n-p} fractional factorial designs using the defining generators. When the alias structure is very complex, as is the case in fractional factorials with 3 or more levels or when partial aliasing is involved, a more general procedure is needed. This process is applicable even for 2^{n-p} fractional designs but is somewhat more complex.

Let $\mathbf{X}$ be the design matrix for a given experiment. Let $\mathbf{X}_1$ be the portion of the design matrix that contains the effects for which the aliases are desired and $\mathbf{X}_2$ contain the columns of $\mathbf{X}$ not contained in $\mathbf{X}_1$. The alias matrix is designed as $\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{X}_2$. The following example will illustrate the process. Suppose we have an experiment with four factors A, B, C, D in a 2^{4-1} fractional factorial design with generator, $I = ABCD$. The design matrix would be

$$\mathbf{X} = \begin{array}{c} \begin{array}{cccccccccccccccc} I & A & B & AB & C & AC & BC & D & AD & BD & ABC & ABD & CD & ACD & BCD & ABCD \end{array} \\ \begin{pmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

This is a resolution IV design with 8 runs so there are 8 alias sets with 2 elements each. The main effects are not confounded with each other or any two-way interaction but the two-way interactions will have some confounding with each other. Thus, we decide to identify what effects are confounded with the main effects A, B, C, D and three two-way interactions, AB, AC, BC. Thus, the necessary $\mathbf{X}_1$ and $\mathbf{X}_2$ matrices are

$$\mathbf{X}_1 = \begin{array}{c} \begin{array}{cccccccc} I & A & B & AB & C & AC & BC & D \end{array} \\ \begin{pmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array} \qquad \mathbf{X}_2 = \begin{array}{c} \begin{array}{cccccccc} AD & BD & ABC & ABD & CD & ACD & BCD & ABCD \end{array} \\ \begin{pmatrix} 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

The alias matrix is obtained from the matrix operations:

$$
\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{X}_2 =
\begin{array}{c c}
\begin{array}{cccccccc} AD & BD & ABC & ABD & CD & ACD & BCD & ABCD \end{array} & \\
\left(\begin{array}{cccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{array}\right) &
\begin{array}{c} I \\ A \\ B \\ AB \\ C \\ AC \\ BC \\ D \end{array}
\end{array}
$$

The alias relations are then obtained from the alias matrix by identifying the effects in the row and column headings associated with each 1 in the matrix:

Effect	Confounded Effect
I	ABCD
A	BCD
B	ACD
AB	CD
C	ABD
AC	BD
BC	AD
D	ABC

Finished @ Approx 27 min mark Monday 4/18/22 (week 13, lecture 33)