

# IE 361 Module 23

## Design and Analysis of Experiments: Part 4

Prof.s Stephen B. Vardeman and Max D. Morris

Reading: Section 7.1, *Statistical Quality Assurance Methods for Engineers*

In this module we discuss what can be done in a many-factor context where the number of possible combinations of levels of  $p$  factors is so large that doing a complete factorial experiment is not practically possible, and one must make do with information from only some fraction of the factorial. We limit our discussion to fractions of  $2^p$  studies, and consider first half fractions of these, and then other fractions that are powers of  $1/2$ .

## **Motivation for the Subject of $2^{p-q}$ Studies and a Qualitative Introduction**

For  $p$  factors, even  $2^p$  gets big fast

$$\text{for } p = 10 \quad 2^p = 1024$$

**Example 23-1** (From an article by Hendrix in 1979 *Chemtech*) A 15-factor chemical experiment had factors

Factor	Levels	Factor	Levels
A-Coating Roll Temp	115 <sup>o</sup> vs 125 <sup>o</sup>	J-Feed Air to Dryer Preheat	Yes vs No
B-Solvent	Recycled vs Refined	K-Dibutylfutile in Formula	12% vs 15%
C-Polymer X-12 Preheat	No vs Yes	L-Surfactant in Formula	.5% vs 1%
D-Web Type	LX-14 vs LB-17	M-Dispersant in Formula	.1% vs .2%
E-Coating Roll Tension	30 vs 40	N-Wetting Agent in Formula	1.5% vs 2.5%
F-Number of Chill Roll	1 vs 2	O-Time Lapse	10min vs 30min
G-Drying Roll Temp	75 <sup>o</sup> vs 80 <sup>o</sup>	P-Mixer Agitation Speed	100rpm vs 250rpm
H-Humidity of Air Feed	75% vs 90%		

and response variable

$y$  = a measure of product cold crack resistance

A full factorial here would require at least  $2^{15} = 32,768$  runs of the process! An obvious "solution" is to collect data for only some (a fraction) of all possible combinations of levels of the factors.

Qualitative points that ought to be "obvious" *apriori* if a fraction of a factorial is used as an experimental design are that

- there must be some information loss (relative to the full factorial)
- some ambiguity must inevitably follow because of the loss
- careful planning and wise analysis are needed to hold these to a minimum

**Example 23-2** (hypothetical)  $2^{2-1}$  ... a half fraction of a  $2 \times 2$  factorial. The full  $2^2$  factorial structure is shown in Figure 1.

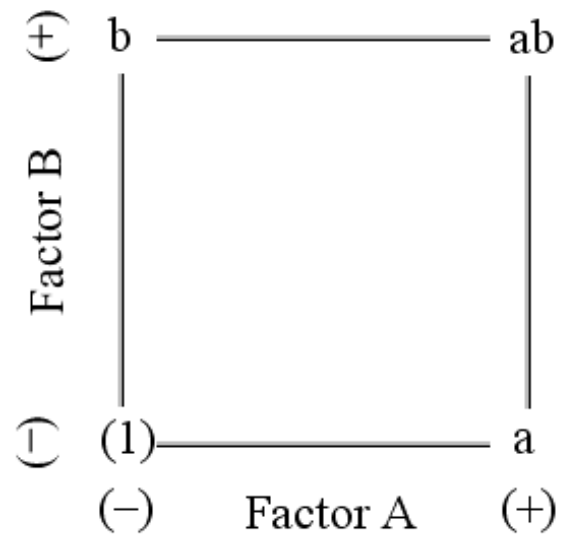


Figure 1: A  $2^2$  Full Factorial Layout

Here, if combination (1) is used, combination ab must be employed or else one learns nothing about the action of one of the factors. (If a is used, then b must be employed or one learns nothing about the action of one of the factors.) On the other hand, if there is a big difference in response between the two combinations included in the half fraction, one doesn't know whether to attribute this to one or the other or to the interaction effect of the factors A and B.

**Example 23-3** (hypothetical) a  $2^{3-1}$  ... Suppose that  $2^3$  factorial effects are

$$\begin{aligned}\mu_{...} &= 10, \alpha_2 = 3, \beta_2 = 1, \gamma_2 = 2, \alpha\beta_{22} = 2, \\ \alpha\gamma_{22} &= 0, \beta\gamma_{22} = 0, \alpha\beta\gamma_{222} = 0\end{aligned}$$

The corresponding means are then as pictured in Figure 2.

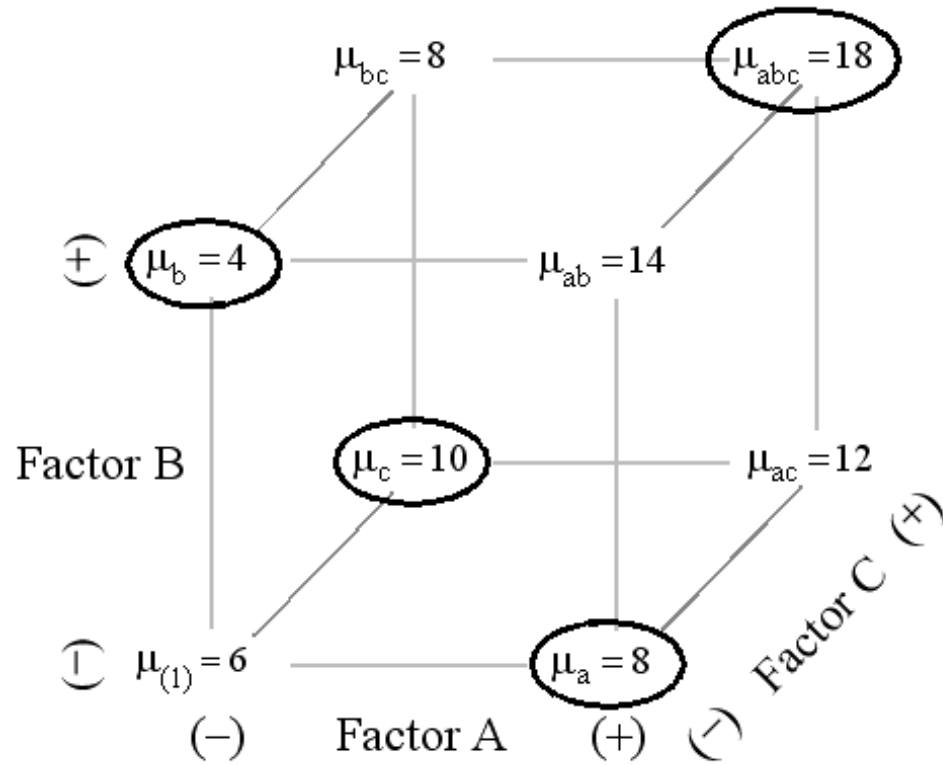


Figure 2: A 2<sup>3</sup> Factorial With a "Good" Half Fraction Indicated With Circled Corners

Suppose further that one gets data adequate to essentially reveal the mean responses for combinations a, b, c and abc (the 4 corners circled in Figure 2) but has no data on the other combinations. (This choice of 4 corners has admirable symmetry. If one collapses the cube in any direction one is left with a complete factorial arrangement in the two other factors. That means that if, in fact, one of the 3 factors is "inert" doing nothing to affect response, one has a full factorial in the 2 factors that matter.) How then might one try to assess the  $2^3$  factorial effects if presented with information from the circled corners on Figure 2?

Remember that with the  $2^3$  as pictured in Figure 2,

$$\alpha_2 = \text{"right face average"} - \text{"grand average"} = 3$$

A "half-fraction version" of this might be

$$\begin{aligned}\alpha_2^* &= \text{"available right face average"} - \text{"available grand average"} \\ &= 13 - 10 \\ &= 3\end{aligned}$$

!!! Here  $\alpha_2^* = \alpha_2$ !!! Is this something for nothing? Can we learn about the A main effect using data from only "4 corners"?

But note that a similar calculation for the C main effect gives

$$\begin{aligned}\gamma_2^* &= \text{"available back face average"} - \text{"available grand average"} \\ &= 14 - 10 \\ &= 4\end{aligned}$$

$$??? \quad 4 = \gamma_2^* \neq \gamma_2 = 2 \quad ???$$

The general story behind this situation is that for this  $2^{3-1}$  fractional factorial

$$\alpha_2^* = \alpha_2 + \beta\gamma_{22} \quad \text{and} \quad \gamma_2^* = \gamma_2 + \alpha\beta_{22}$$

We were able to "recover"  $\alpha_2$  using only  $\alpha_2^*$  (that is based on only half of the corners of the cube) because  $\beta\gamma_{22} = 0$ . We were unable to "recover"  $\gamma_2$  using only  $\gamma_2^*$  (that is based on only half of the corners of the cube) because  $\alpha\beta_{22} = 2 \neq 0$ . We can really only know  $\alpha_2 + \beta\gamma_{22}$  (and not  $\alpha_2$ ) and  $\gamma_2 + \alpha\beta_{22}$  (and not  $\gamma_2$ ) based on the half fraction. This is an example of confounding/aliasing/ambiguity that of necessity comes with use of only a fractional factorial.

## Half Fractions of $2^p$ Factorial Studies

The issues to be addressed in order to use  $2^{p-q}$  fractional factorials are:

- how to rationally choose  $\frac{1}{2^q}$  out of  $2^p$  combinations for study
- how to determine the corresponding aliasing/confounding pattern
- how to do data analysis

We proceed to consider these questions first in the context of half fractions (the  $q = 1$  case) then for general  $1/2^q$  fractions.

## **Choice of standard half fractions of $2^p$ factorials**

To choose what we will call a "standard" (a "good") half fraction of a  $2^p$  factorial, we will

- write out signs for specifying levels for all possible combinations of levels of the “first”  $p - 1$  factors
- then “multiply” these together for a given combination of the “first” factors to arrive at a corresponding level to use for the “last” factor

**Example 23-4** (A  $2^{4-1}$  fractional factorial) With 4 two-level factors A, B, C and D, one proceeds as in the following table. (One multiplies  $-$  and  $+$  signs as if they were  $\pm 1$ 's and the last column gives the indicated combination of factor levels for the row, using the special  $2^p$  naming convention introduced in Module 22.

A	B	C	Product (used for D)	Combination
–	–	–	–	(1)
+	–	–	+	ad
–	+	–	+	bd
+	+	–	–	ab
–	–	+	+	cd
+	–	+	–	ac
–	+	+	–	bc
+	+	+	+	abcd

**Example 23-5** R. Snee in a 1985 *ASQC Technical Supplement* discussed a  $2^{5-1}$  chemical process study. The factors and their levels were as in the following table.

Factor	—	vs	+
A-Solvent/Reactant	low	vs	high
B-Catalyst/Reactant	.025	vs	.035
C-Temperature	150°	vs	160°
D-Reactant Purity	92%	vs	96%
E-pH of Reactant	8.0	vs	8.7

In Snee's study, the response variable was

$$y = \text{color index}$$

Snee's (unreplicated) data were

combination	$y$	combination	$y$
e	−.63	d	6.79
a	2.51	ade	6.47
b	−2.68	bde	3.45
abe	−1.66	abd	5.68
c	2.06	cde	5.22
ace	1.22	acd	9.38
bce	−2.09	bcd	4.30
abc	1.93	abcde	4.05

These are data from half of all 32 combinations of 2 levels of each of the 5 factors (half of all possible labels of combinations based on the 5 letters a,b,c,d and e are given above, namely those involving an odd number of letters). In fact, Snee followed the standard recommendation for choosing this half fraction.

## Determining the “alias structure” of a half fraction (the implied pattern of ambiguities)

To understand the pattern of ambiguities one is left with upon using a standard half fraction of a  $2^p$  factorial, we will use a method of formal multiplication, beginning from a so-called “generator” that represents the way in which the half fraction was chosen. The **generator** is of the form

name of “last” factor  $\leftrightarrow$  product of names of “first” factors

The rules of formal multiplication are that

- any letter  $\times$  I  $\leftrightarrow$  the same letter
- letter  $\times$  same letter  $\leftrightarrow$  I

**Example 23-3 (continued)** (the  $2^{3-1}$  numerical example used above) The generator here is

$$C \leftrightarrow AB$$

We can multiply through by C to obtain the so called “**defining relation**”

$$I \leftrightarrow ABC$$

This first says that the ABC 3-factor interaction  $\alpha\beta\gamma_{222}$  is aliased with the grand mean. That is, only

$$\mu \dots + \alpha\beta\gamma_{222}$$

can be evaluated based on the half-fraction, not  $\alpha\beta\gamma_{222}$  alone.

Multiplying through the defining relation by any set of letters of interest produces a statement of what effect(s) are “aliased with” the corresponding effect. For example, we see that

$$A \leftrightarrow BC$$

(read “the A main effect is aliased with the BC 2-factor interaction). Similarly

$$C \leftrightarrow AB$$

as was illustrated earlier. In fact, the whole alias structure is

$$I \leftrightarrow ABC$$

$$A \leftrightarrow BC$$

$$B \leftrightarrow AC$$

$$C \leftrightarrow AB$$

and we see that  $2^3$  factorial effects are aliased in 4 pairs.

The technical meaning of aliasing is that only sums of effects can be learned from the half fraction study, not individual effects.

**Example 23-4 continued** (the  $2^{4-1}$  again) With the generator

$$D \leftrightarrow ABC$$

the defining relation is

$$I \leftrightarrow ABCD$$

From this, e.g., we see that the AB 2-factor interaction is aliased with the CD 2-factor interaction.

**Example 23-5 continued** Snee's  $2^{5-1}$  study had generator

$$E \leftrightarrow ABCD$$

and hence defining relation

$$I \leftrightarrow ABCDE$$

From this one sees, e.g., that the AB 2-factor interaction is aliased with the CDE 3-factor interaction.

## Data analysis for standard half fractions

To do data analysis for a half fraction of a  $2^p$  study, one may

- initially temporarily ignore the “last” factor, treat the data as a full factorial in the “first”  $p - 1$  factors, and judge the statistical significance and practical importance of estimates derived from the Yates algorithm
- then interpret these estimates in light of the alias structure as estimates of appropriate sums of  $2^p$  effects.

Where there is some replication (not all  $2^{p-1}$  sample sizes are 1) confidence intervals can be made for the (sums of) effects using

$$\hat{E} \pm t s_{\text{pooled}} \frac{1}{2^{p-1}} \sqrt{\sum \frac{1}{n_{\text{comb}}}}$$

where

$$s_{\text{pooled}}^2 = \frac{\sum \left( (n_{\text{combination}} - 1) s_{\text{combination}}^2 \right)}{\sum (n_{\text{combination}} - 1)}$$

and the appropriate degrees of freedom for  $t$  are

$$\sum (n_{\text{combination}} - 1) = n - 2^{p-1}$$

Lacking any replication, normal plotting of the output of the Yates algorithm (ignoring the “last” factor) can be used in judging statistical significance.

**Example 23-6** (another hypothetical  $2^{3-1}$ ) Suppose

$$n_a = 1, y_a = 5, n_b = 2, \bar{y}_b = 3, s_b^2 = 1.5, n_c = 1, y_c = 2.5, \text{ and}$$

$$n_{abc} = 3, \bar{y}_{abc} = 5.5, s_{abc}^2 = 1.8$$

Then listing the 4 combinations in Yates standard order as regards the "first" 2 factors (i.e. ignoring the "last" factor), the ( $p - 1 = 2$  cycle) Yates algorithm is applied to

comb	$\bar{y}$
c	2.5
a	5.0
b	3.0
abc	3.0

Confidence intervals based on the output of the algorithm would be made using

$$s_{\text{pooled}}^2 = \frac{0 + (2 - 1)1.5 + 0 + (3 - 1)1.8}{0 + (2 - 1) + 0 + (3 - 1)}$$

These have the form

$$\hat{E} \pm t_{3s_{\text{pooled}}} \frac{1}{2^{3-1}} \sqrt{\frac{1}{1} + \frac{1}{2} + \frac{1}{1} + \frac{1}{3}}$$

**Example 23-5 continued** Snee's  $2^{5-1}$  study had no replication. Ignoring factor E temporarily, the (4-cycle) Yates algorithm can be applied to the 16 responses exactly as listed earlier (they are in Yates order as regards the first 4 factors). The result is the set of estimates below.

combination	$y$	estimate (16 divisor)
e	−.63	2.875
a	2.51	.823
b	−2.68	−1.253
abe	−1.66	.055
c	2.06	.384
ace	1.22	.064
bce	−2.09	.041
abc	1.93	.001
d	6.79	2.793
ade	6.47	−.095
bde	3.45	−.045
abd	5.68	−.288
cde	5.22	−.314
acd	9.38	.186
bcd	4.30	−.306
abcde	4.05	−.871

A normal plot of the (last 15) estimates is in Figure 3 and suggests that at most 4 sums of effects are distinguishable from background variation.

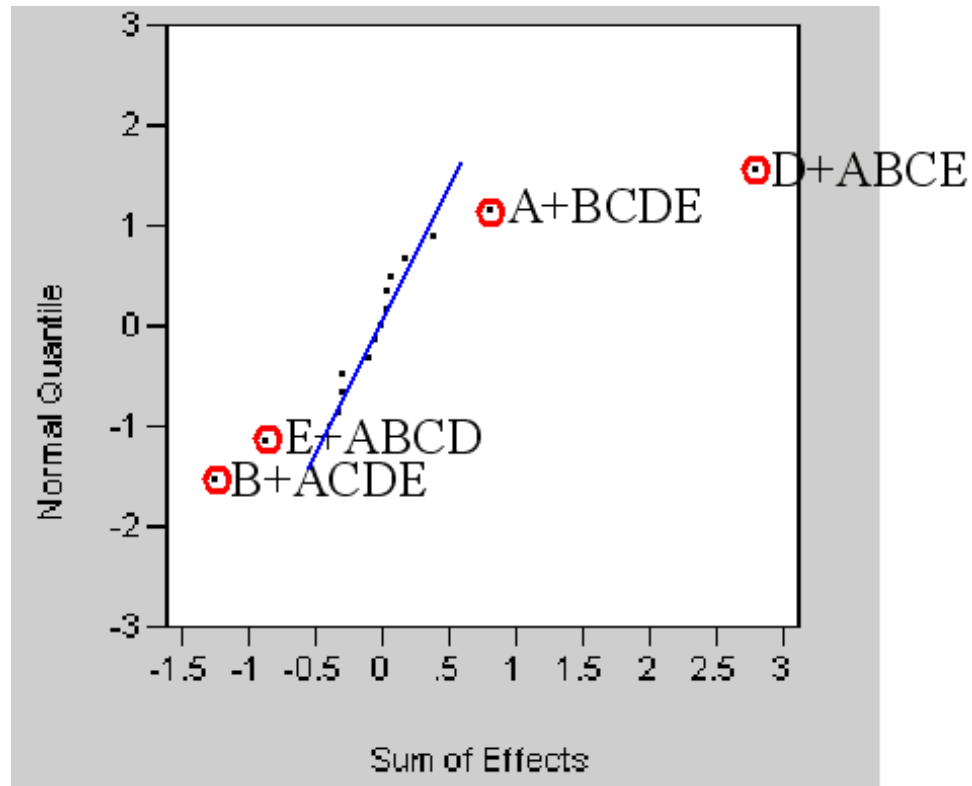


Figure 3: A Normal Plot of 15 Fitted Sums of Effects From Snee's  $2^{5-1}$  Study

Tentative engineering conclusions based on Snee's study were that for uniform color index, attention must be paid to controlling/reducing variation in the following (in decreasing order of importance):

- Factor D, Reactant Purity
- Factor B, Catalyst/Reactant Ratio
- Factor E, pH of Reactant
- Factor A, Solvent/Reactant Ratio

# $\frac{1}{2^q}$ Fractions of $2^p$ Factorial Studies

The issues in the use of smaller than half fractions of  $2^p$  factorials are completely parallel to those for half fractions, namely

- how to rationally choose  $\frac{1}{2^q}$  of  $2^p$  possible combinations of levels of  $p$  2-level factors
- how to determine the corresponding aliasing/confounding pattern
- how to do data analysis

The answers here are the natural generalizations of the half fraction answers just discussed.

## Choice of $\frac{1}{2^q}$ fractions of $2^p$ factorials

To choose a  $1/2^q$  fraction of a  $2^p$  factorial, we will

- write out signs for specifying levels for all possible combinations of levels of the “first”  $p - q$  factors
- pick  $q$  different groups of the first  $p - q$  factors and use the products of the signs corresponding to members of the groups to specify levels for the “last”  $q$  factors

**Example 23-7** Hanson and Best in a presentation at the 1986 annual meeting of the American Statistical Association reported on an experiment for the

development of a catalyst for producing ethyleneamines by the amination of monoethanolamine involving  $p = 5$  factors. These factors and their levels were

Factor	—		+
A-Ni/Re Ratio	2/1	vs	20/1
B-Precipitant	(NH <sub>4</sub> ) <sub>2</sub> CO	vs	none
C-Calcining Temp	300°	vs	500°
D-Reduction Temp	300°	vs	500°
E-Support Used	alpha-alumina	vs	silica alumina

The response of interest was

$$y = \% \text{ water produced}$$

The investigators decided against a full factorial, choosing instead to use a  $2^{5-2}$  design. That is, they chose to use  $q = 2$  and a  $\frac{1}{4}$  fraction of the full

$2^5$  study. They ran  $2^{5-2} = 8$  out of the  $2^5 = 32$  possible A, B, C, D, E combinations

The (somewhat arbitrary) choice was made to use ABC sign products to choose levels of D, and BC sign products to choose levels of E. (Other choices are possible and lead to different aliasing patterns that might for some other studies be preferred by the engineer in charge.) The choice used is summarized in the following table, whose last column specifies the 8 combinations of levels of the 5 factors used in the study in the special  $2^p$  factorial notation.

A	B	C	ABC Product (for D)	BC Product (for E)	Combination
-	-	-	-	+	e
+	-	-	+	+	ade
-	+	-	+	-	bd
+	+	-	-	-	ab
-	-	+	+	-	cd
+	-	+	-	-	ac
-	+	+	-	+	bce
+	+	+	+	+	abcde

The data obtained were as listed and summarized below.

combination	$y$	$\bar{y}$	$s^2$
e	8.70, 11.60, 9.00	9.767	2.543
ade	26.80	26.800	
bd	24.88	24.880	
ab	33.15	33.150	
cd	28.90, 30.98	29.940	2.163
ac	30.20	30.200	
bce	8.00, 8.69	8.345	.238
abcde	29.30	29.300	

## Determining the “alias structure” of a $\frac{1}{2^q}$ fraction of a $2^p$ factorial

To understand the pattern of ambiguities one faces with a particular choice of a  $\frac{1}{2^q}$  fraction of a  $2^p$  factorial, we will use the formal multiplication, beginning from the  $q$  **generators** that represent the way in which the  $\frac{1}{2^q}$  fraction was chosen. To find the **defining relation** (the list of all products "equivalent to" I) we first convert the generators to statements of products equivalent to I, and then multiply these in pairs, then in triples, then in sets of four, etc. The letter I will have  $2^q - 1$  equivalent products ... i.e. the  $2^p$  factorial effects are aliased in  $2^{p-q}$  different groups of  $2^q$  each.

**Example 23-7 continued** In the catalyst example, the generators were

$$D \leftrightarrow ABC \quad \text{and} \quad E \leftrightarrow BC$$

so

$$I \leftrightarrow ABCD \text{ and } I \leftrightarrow BCE$$

Further, multiplying these two we get

$$I \cdot I \leftrightarrow (ABCD) \cdot (BCE) \text{ i.e. } I \leftrightarrow ADE$$

So the whole defining relation for the catalyst study is

$$I \leftrightarrow ABCD \leftrightarrow BCE \leftrightarrow ADE$$

and therefore effects are aliased in 8 groups of 4. For example, multiplying through the defining relation by A gives

$$A \leftrightarrow BCD \leftrightarrow ABCE \leftrightarrow DE$$

and we see that the A main effect is aliased with the DE 2-factor interaction (among other things).

## Data analysis for standard $2^{p-q}$ studies

To do data analysis for a  $2^{p-q}$  study, one may

- initially temporarily ignore the “last”  $q$  factors, and treating the data as a full factorial in the “first”  $p - q$  factors, judge the statistical significance and practical importance of estimates produced by the Yates algorithm
- then interpret these estimates in light of the alias structure as estimates of appropriate sums of  $2^p$  effects.

Where there is some replication (not all  $2^{p-q}$  sample sizes are 1) confidence intervals can be made for the (sums of) effects using

$$\hat{E} \pm ts_{\text{pooled}} \frac{1}{2^{p-q}} \sqrt{\sum \frac{1}{n_{\text{comb}}}}$$

where

$$s_{\text{pooled}}^2 = \frac{\sum \left( (n_{\text{combination}} - 1) s_{\text{combination}}^2 \right)}{\sum (n_{\text{combination}} - 1)}$$

and the appropriate degrees of freedom for  $t$  are

$$\sum (n_{\text{combination}} - 1) = n - 2^{p-q}$$

Lacking any replication, normal plotting of the output of the Yates algorithm (ignoring the “last” factor) can be used in judging statistical significance.

**Example 23-7 continued** In the catalyst example the 8 sample means,  $\bar{y}$ , listed before were in Yates standard order for factors A, B and C (the “first”  $p - q = 3$ ) ignoring D and E (the “last”  $q = 2$ ). So the ( $p - q = 3$  cycle) Yates algorithm can be applied to them in the order listed. The following table shows the first two and last columns of the Yates table and then records what the estimates produced by the algorithm attempt to approximate.

combination	$\bar{y}$	estimate	sum estimated
e	9.767	24.048	grand mean+aliases
ade	26.800	5.815	A main effect+aliases
bd	24.880	-.129	B main effect+aliases
ab	33.150	1.492	AB interaction+aliases
cd	29.940	.399	C main effect+aliases
ac	30.200	-.511	AC interaction+aliases
bce	8.345	-5.495	BC interaction+aliases
abcde	29.300	3.682	ABC interaction+aliases

Since the original data had 3 sample sizes larger than 1, statistical significance/detectability of these can be judged using confidence limits for sums of effects. First,

$$s_{\text{pooled}}^2 = \frac{(3 - 1)(2.543) + (2 - 1)(2.163) + (2 - 1)(.238)}{(3 - 1) + (2 - 1) + (2 - 1)} = 1.872$$

So  $s_{\text{pooled}} = \sqrt{1.872} = 1.368$ , and this can be used as a measure of background noise and as a basic ingredient of confidence intervals for the sums of effects.

$s_{\text{pooled}}$  has 4 associated degrees of freedom. So if, e.g., 95% confidence intervals for the sums of effects are desired, the “+/- part” of the confidence interval formula becomes

$$\pm 2.776(1.368) \frac{1}{2^3} \sqrt{\frac{1}{3} + \frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{1} + \frac{1}{2} + \frac{1}{1}}$$

i.e.  $\pm 1.195$

We therefore might judge any estimate larger in absolute value than 1.195 to represent a sum of effects clearly large enough to see above the background experimental variation.

Then the “detectable” sums are (in order of magnitude):

sum	estimate
$\alpha_2 + \beta\gamma\delta_{222} + \alpha\beta\gamma\epsilon_{2222} + \delta\epsilon_{22}$	5.815
$\beta\gamma_{22} + \alpha\delta_{22} + \epsilon_2 + \alpha\beta\gamma\delta\epsilon_{22222}$	-5.495
$\alpha\beta\gamma_{222} + \delta_2 + \alpha\epsilon_{22} + \beta\gamma\delta\epsilon_{2222}$	3.682
$\alpha\beta_{22} + \gamma\delta_{22} + \alpha\gamma\epsilon_{222} + \beta\delta\epsilon_{222}$	1.492

Happily, the last of these is smaller in magnitude than the other 3, but there are at least 4 *a priori* equally plausible interpretations of the possibility that these 3 are really driven by a single effect each ... there might be important

- A main effects, E main effects, and D Main effects
- A main effects, E main effects, and AE 2-factor interactions
- A main effects, AD 2-factor interactions, and D main effects
- DE 2-factor interactions, E main effects, D main effects

In fact, a follow-up study confirmed the importance of the D main effect (and made the first of these most attractive).

If the A (Ni/Re ratio) main effect, the E (Support Type) main effect and the D (Reduction Temp) main effect are indeed the most important determiners

of  $y$ , and large  $y$  is desirable, the signs of the estimates indicate the need for “high A” (20/1 Ni/Re ratio), “low E” (alpha-alumina support) and “high D” (500° reduction temp).

Notice !!!! The larger  $q$ , the larger the inevitable ambiguity of interpretation of the fractional factorial results and the more likely the need for follow-up study. *Small fractions are really most useful as screening studies, to pick a few likely candidates out of many potentially important factors for subsequent more detailed study.*

We end with an extreme example of large  $q$ , i.e. a small fraction.

**Example 23-1 continued** The Hendrix chemical process study involved  $p = 15$  factors A, B, C, D, E, F, G, H, J, K, L, M, N, O, P

(the factor names and levels were given earlier).

$p - q = 4$ , i.e., only  $2^4 = 16$  combinations were run !!!!! This was a  $\frac{1}{2^{15-4}} = \frac{1}{2048}$  fraction !!!!

The 11 generators used were:

E ↔ ABCD, F ↔ BCD, G ↔ ACD, H ↔ ABC, J ↔ ABD,  
K ↔ CD, L ↔ BD, M ↔ AD, N ↔ BC, O ↔ AC, P ↔ AB

These led to the 16 combinations and (ultimately) the data below.

combination	$y$
eklmnop	14.8
aghjklm	16.3
bfhjkmo	23.5
abefgkp	23.9
cfghlmp	19.6
acefjlo	18.6
bcegjmn	22.3
abchnop	22.2
dfgjnop	17.8
adefhmn	18.9
bdeghlo	23.1
abdjlm	21.8
cdehjpk	16.6
acdgmko	16.7
bcdfkln	23.5
abcdefghijklmnop	24.9

Pretty clearly it isn't sensible to write out the whole defining relation here ... effects are going to be aliased in 16 groups of  $2^{11} = 2048$  effects. But for a most tentative interpretation, let's see what we might glean if the physical system is so simple that only main effects dominate. (Whether this is physically reasonable is a question that needs to be answered by a process expert.)

The 16 observations are listed in Yates order for factors A,B,C and D (ignoring the rest). We therefore begin by running them through the Yates algorithm, with the results below.

combination	$y$	estimate (16 divisor)	sum estimated
eklmnop	14.8	20.28	grand mean + ...
aghjkl	16.3	.13	A + ...
bfhjkmo	23.5	2.87	B + ...
abefgkp	23.9	-.08	AB + P + ...
cfghlmp	19.6	.27	C + ...
acefjlo	18.6	-.08	AC + O + ...
bcegjmn	22.3	-.19	BC + N + ...
abchnop	22.2	.36	ABC + H + ...
dfgjnop	17.8	.13	D + ...
adefhmn	18.9	.03	AD + M + ...
bdeghlo	23.1	.04	BD + L + ...
abdjlm	21.8	-.06	ABD + J + ...
cdehjpk	16.6	-.26	CD + K + ...
acdgkmo	16.7	.29	ACD + G + ...
bcdfkln	23.5	1.06	BCD + F + ...
abcdefghijklmnop	24.9	.11	ABCD + E + ...

There is no replication in this data set, so we're driven to normal plotting in order to judge statistical significance of these estimates. A normal plot of the (last 15) estimates is in Figure 4.

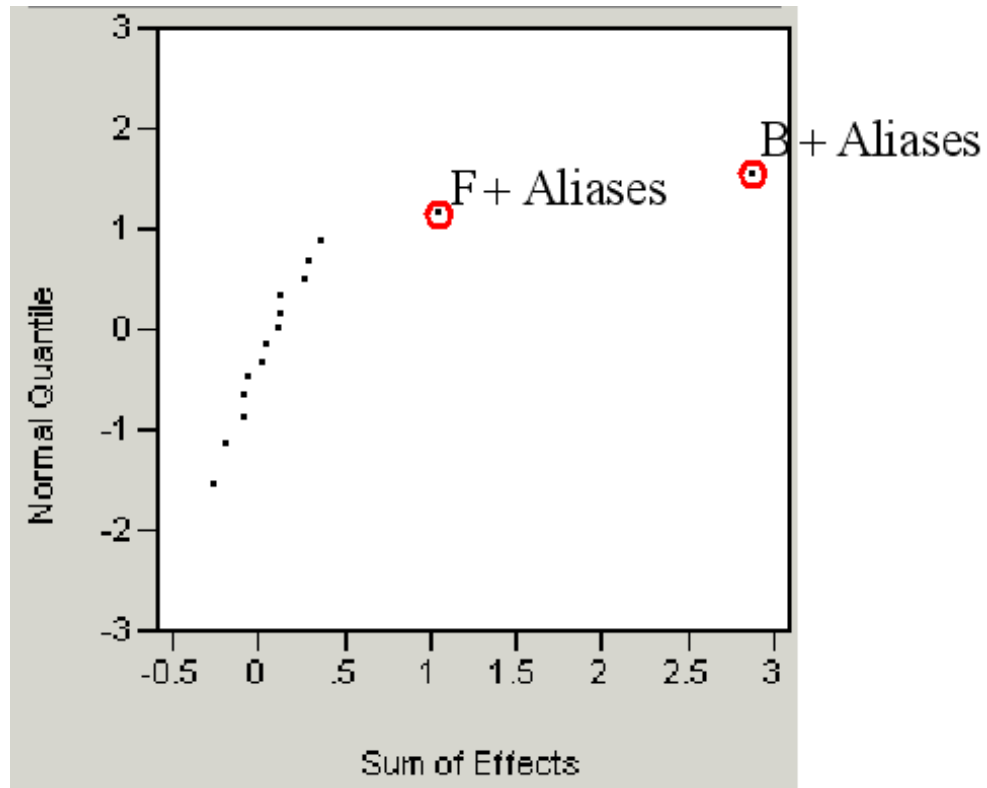


Figure 4: Normal Plot of 15 Estimated Sums of Effects from the  $2^{15-11}$  Chemical Process Study

The plot shows that there are two fitted sums of effects that are clearly statistically detectable. As each of these stands for a sum of 2048 effects, no solid final conclusions can be made. But if one assumes that the "big" sums are primarily driven by the main effects appearing in them, a plausible tentative interpretation is that the most important factors appear to be

B (Solvent) and F (# of Chill Rolls)

and for large cold crack resistance "high B" (refined solvent) and "high F" (2 chill rolls) appear best.

Note that the analysis does point out what is in retrospect quite obvious, namely that it is those combinations in the data set with "high B" and "high F" that have the largest  $y$ 's.