Two Level Fractional Factorials

Design of Experiments - Montgomery
Sections 8-1 – 8-3

Fractional Factorials

- May not have sources for complete factorial design
- Number of runs required for factorial grows quickly
  - Consider 2^k design
  - If $k = 7 \rightarrow 128$ runs required
  - Can estimate 127 effects
  - Only 7 df for main effects
  - 99 df are for interactions of order ≥ 3
- Often system driven by low order effects
- Would like design
  - To utilize this sparsity principle
  - Can be projected into larger designs
- Fractional factorials provide these options

Example

- Suppose you were designing a new car for mileage
- Wanted to consider (2 levels each)
  - Engine Size
  - Number of cylinders
  - Drag
  - Weight
  - Automatic vs Manual
  - Shape
  - Tires
  - Suspension
  - Gas Tank Size
- Only have resources for 2^7 design

  If you drop two factors for a complete factorial, could discard significant main effects or lower order interactions

  Want option to keep all nine factors in model

  Must assume higher order interactions insignificant

  Are six and seven order interactions meaningful?

Two-Level Fractional Factorials

- Assume certain higher order interactions negligible
- Can then collect more info on lower level effects
- Example: Latin Square (2^3 factorial)
  - Let A=Block Factor 1, B=Block Factor 2, and C=Treatment

A	B	C	AB	AC	BC	ABC	Symbol
-	-	-	+	+	+	-	(1)
+	-	-	-	-	+	+	a
-	+	-	-	+	-	+	b
+	+	-	+	-	-	-	ab
-	-	+	+	-	-	+	c
+	-	+	-	+	-	-	ac
-	+	+	-	-	+	-	bc
+	+	+	+	+	+	+	abc

  - Utilize only four observations instead of eight

A \ B	−	+
−	+	−
+	−	+

A \ B	−	+
−	−	+
+	+	−

  - Observed combinations associated with ABC column

1 May observe C, B, A, ABC

2 May observe (1), BC, AC, and AB

- Since only four obs, certain factors confounded
- Use association column to determine confounding
- Association column known as the generator
- Use generator in the following manner

  I=ABC → A=BC, B=AC, and C=AB

- Thus if additivity, can estimate A, B, and C

Assumptions and Expectations

- Used when

  Runs expensive and variance estimates available

  Screening experiments when many factors considered

  Sequential design analysis possible (put fractions together)

- Interest in main effects and low order interactions
- Prepared to assume certain interactions are zero
- Not primarily interested in estimate of variance

Emphasis is on finding small experiments in which a high percentage of df are used for estimation of low order effects

- Notation
  - Full factorial is 2^k
  - Fractional Factorial is 2^{k-p}
  - Degree of fraction is 2^{-p}

Half-Fraction 2^k Factorials

- This is one half the usual number of runs
- Similar to blocking procedure
  - Choose a **generator** which divides effects into two
  - Based on pluses and minuses of one factor
  - **Defining Relation**: $I =$ generator
- Consider three factor but only 4 runs possible

A	B	C	AB	AC	BC	ABC	Symbol
−	−	−	+	+	+	−	(1)
+	−	−	−	−	+	+	a
−	+	−	−	+	−	+	b
+	+	−	+	−	−	−	ab
−	−	+	+	−	−	+	c
+	−	+	−	+	−	−	ac
−	+	+	−	−	+	−	bc
+	+	+	+	+	+	+	abc

- Select I=ABC, get groups (a,b,c,abc) and (ab,ac,bc,(1))
- Select I=A, get groups (a,ab,ac,abc) and ((1),b,c,bc)
- Use generator to determine confounded effects

Half-Fraction 2^k Factorials

- Consider I=ABC and the group (a,b,c,abc)

A	B	C	AB	AC	BC	ABC	Symbol
−	−	−	+	+	+	−	(1)
+	−	−	−	−	+	+	a
−	+	−	−	+	−	+	b
+	+	−	+	−	−	−	ab
−	−	+	+	−	−	+	c
+	−	+	−	+	−	−	ac
−	+	+	−	−	+	−	bc
+	+	+	+	+	+	+	abc

- Each effect is estimated as follows

$$l_A = .5(a-b-c+abc) \quad l_{BC} = .5(a-b-c+abc)$$
$$l_B = .5(-a+b-c+abc) \quad l_{AC} = .5(-a+b-c+abc)$$
$$l_C = .5(-a-b+c+abc) \quad l_{AB} = .5(-a-b+c+abc)$$

- Cannot differentiate between

  A and BC, B and AC, and C and AB

- Can use defining relation to determine confounding
- "Multiply" each side by an effect

  (A)I = A = (A)ABC = A^2BC=BC

  (AB)I = AB = (AB)ABC = A^2B^2C=C

- Linear combinations estimate

  A+BC, B+AC, and C+AB

Half-Fraction 2^k Factorials

- Suppose we use other grouping ($I = -ABC$)

$$l_A = .5(ab + ac - bc - (1)) \quad l_{BC} = .5(-ab - ac + bc + (1))$$
$$l_B = .5(ab - ac + bc - (1)) \quad l_{AC} = .5(-ab - ac + bc + (1))$$
$$l_C = .5(-ab + ac + bc - (1)) \quad l_{AB} = .5(ab - ac - bc + (1))$$

- $l_A = -l_{BC}$ so combination is estimating $A - BC$
- If both fractions were run
  - Could separately estimate effects
  - ABC is confounded with blocks
- Similar to factorial in two blocks

Can piece together fractional factorial into bigger design if appropriate

Another Example

Consider a 2^{5-1} fractional factorial. Usually would have 32 runs, but we will have 16. Let us use the defining relation $I = ABCDE$. This means we will have the following confounded effects

A and BCDE	AB and CDE	ACD and BE
AC and BDE	AD and BCE	ACE and BD
AE and BCD	ABC and DE	ABCD and E
ABD and CE	ABE and CD	ABDE and C
ADE and BC	ABCE and D	ACDE and B

Main effects confounded with 3rd order interactions, 1st order interactions confounded with 2nd order interactions

Known as a Resolution V Design

Using Yates' Algorithm

- For 2^{k-1} design, will set up $k-1$ columns
- Select $k-1$ factors and write in standard order
- Multiply columns to determine sign of last factor
- See Table 8-5 for 2^{5-1} example
- Consider these results from previous example

Combination	y	1	2	3	4	Effect
e	15	25	55	110	185	
a	10	30	55	75	35	A+BCDE
b	5	40	30	20	-35	B+ACDE
abe	25	15	45	15	15	AB+CDE
c	15	15	15	-20	15	C+ABDE
ace	25	15	5	-15	-15	AC+BDE
bce	10	30	10	10	-45	BC+ADE
abc	5	15	5	5	-35	ABC+DE
d	5	-5	5	0	-35	D+ABCE
ade	10	20	-25	15	-5	AD+BCE
bde	5	10	0	-10	5	BD+ACE
abd	10	-5	-15	-5	-5	ABD+CE
cde	15	5	25	-30	15	CD+ABE
acd	15	5	-15	-15	5	ACD+BE
bcd	5	0	0	-40	15	BCD+AE
abcde	10	5	5	5	45	ABCD+E

Resolution

- A design of resolution R is one in which no p-factor effect is confounded with any other effect containing less than $R - p$ factors
  - Resolution III design does not confound main effects with other main effects.
  - Resolution IV design does not confound main effects with two-factor interactions but does confound two-factor interactions with other two-factor interactions.
  - Resolution V design does not confound any main effects and two factor interactions with each other.
  - Resolution III - can estimate main effects if you assume no interaction
  - Resolution IV - can estimate main effects without assuming two-factor interactions negligible
  - Resolution V - can estimate main and two-factor interactions if assume higher order terms negligible.
- Resolution also defined by length of shortest word in defining relation
  - I=ABC is Resolution III
  - I=ABCD is Resolution IV
  - I=ABCDE is Resolution V

Construction of a 2^{k-1} fractional factorial with highest resolution

1. Write a full factorial design for the first $k-1$ variables
2. Associate the kth variable with $\pm$ interaction of $k-1$

- Consider 2^{4-1} design

A	B	C	$D = ABC$	Effect	Combination
$-$	$-$	$-$	$-$	(1)	(1)
$+$	$-$	$-$	$+$	a	ad
$-$	$+$	$-$	$+$	b	bd
$+$	$+$	$-$	$-$	ab	ab
$-$	$-$	$+$	$+$	c	cd
$+$	$-$	$+$	$-$	ac	ac
$-$	$+$	$+$	$-$	bc	bc
$+$	$+$	$+$	$+$	abc	abcd

- $I = ABCD \rightarrow$ Resolution 4 design

Alternative View of Half Fraction Design

- Consider any 2^{k-1} design
- If we collapse design by omitting variable
- Remaining is a 2^{k^*} full factorial ($k^* = k-1$)
- Can be shown for resolution R, complete factorial for $R-1$ factors
  - Consider 2^{3-1} design (Resolution III)
  - Can view combinations on cube
  - Regardless of direction, you can squeeze cube into square with observation at each corner

Another Example

- Consider the following 2^3 design

Combination	y	1	2	3	Estimate
(1)	5	17	43	81	
a	12	26	38	21	5.25
b	10	13	13	21	5.25
ab	16	25	8	-3	-0.75
c	4	7	9	-5	-1.25
ac	9	6	12	-5	-1.25
bc	11	5	-1	3	0.75
abc	14	3	-2	-1	-0.25

- Consider instead a half fraction with $I = ABC$

A	B	$C = AB$	Combination	y	1	2	Estimate	Effect
$-$	$-$	$+$	c	4	16	40		
$+$	$-$	$-$	a	12	24	12	6.0	$A + BC$
$-$	$+$	$-$	b	10	8	8	4.0	$B + AC$
$+$	$+$	$+$	abc	14	4	-4	-2.0	$C + AB$

- Notice effects are sums of previous estimates
- Could ignore C and this becomes full factorial

Sequential Use of Factorial Designs

- Often more efficient to look at half fraction
- Analyze results
- Decide on best set of runs for next experiment
  - Can add or remove factors
  - Change responses
  - Vary factors over new ranges
- If ambiguities, can run remaining half of factorial
- Only lose estimate of highest order interaction
- Important to always randomize order of runs
- Can use Proc Factex to generate design

The General 2^{k-p} Fractional Factorial

- Must select p independent generators
- Want to have best alias relationships
- Defining relation based on $2^p - 1$ effects
- Often try to maximize the resolution
- Each effect has $2^p - 1$ aliases
- Often assume higher order interactions zero
- Simplifies alias structure
- Table 8-14 summarizes potential generators
  - $k \leq 15$ and $n \leq 128$
  - Results in highest possible resolution
- Table XII : alias relationships for designs with $n \leq 64$
- Can use Yates' in similar fashion to obtain estimates

Example

Consider 2^{5-2} which consists of eight runs. Suppose we choose $I = ABC$ and $I = BDE$. The defining relation is $I = ABC = BDE = ACDE$ so this is a resolution III design. The factor A is aliased with BC, $ABDE$, and CDE.

Consider 2^{11-4} which has 128 runs. We start with a complete 2^7. Consider the factors A, B, C, D, E, G, J with remaining factors F, H, K, L. Define four generators as

- $F = ABCDE$
- $K = ABFJ$
- $L = AEFGK$
- $H = ACEL$

The defining relation is $I = ABCDEF = ABFJK = AEFGKL = ACEHL$. If multiply these together in pairs, we get $I = CDEJK = BCDGKL = BDFHL = BEGJL = BCEFJHKL = CFGHK$. If we multiply these in triples, we get $I = ABCGHJ = ABDEGHK = ACDFG = ADHJL$ and if we multiply all four together, we get $I = DEFGHJ$. Since the shortest word in relation is of length five, this has resolution V.

The General 2^{k-p} Fractional Factorial

- The 2^{k-p} collapses into either a
  - Full factorial
  - Fractional factorial of subset $r \leq k - p$
- Can block fractional factorials if necessary
  - Presented in Table XII
  - Minimum block size for designs is of size 8
  - Block to confound high order interaction
- Blocking may change resolution of design

Resolution III Designs

- Can use this design to efficiently investigate numerous factors
- Can use resolution III to investigate $N - 1$ factors in N runs
- N must be a multiple of 4
  - a 4 runs to investigate 3 factors $\rightarrow 2^{3-1}$ design
  - b 8 runs to investigate 7 factors $\rightarrow 2^{7-4}$ design
  - c 16 runs to investigate 15 factors $\rightarrow 2^{15-11}$ design
- For 2^{3-1} design
  - Each main effect aliased with one two factor interaction
  - Introduced fractional factorials with this design
- For 2^{7-4} design
  - Each main effect aliased with three two factor interactions
  - Often ignore interactions of order ≥ 3
- For 2^{15-11} design
  - Each main effect aliased with seven two factor interactions

Sequential Assembly of Fractions

- Consider 2^{7-4} design with $D = AB$, $E = AC$, $F = BC$, and $G = ABC$
- If factor D important and don't want it confounded
  - Use same generators except $D = -AB$
  - If all three factor interactions zero, can estimate D
  - Can also estimate all interactions concerning D
- Can use two 2^{7-4} to get resolution IV
  - Instead of flipping one sign, flip sign of all factors
  - Generators of even size flip sign
  - Known as folding over
  - Breaks link between main effects and two-factor interactions
- $D = AB$, $E = AC$, $F = BC$, and $G = ABC$
- $D = -AB$, $E = -AC$, $F = -BC$, $G = ABC$

Example

Consider the 2^{3-1} design with $C = AB \to I = ABC$. This is the trivial example because folding over creates the full factorial.

A	B	$C = AB$	Combination	y	1	2	Effect
$-$	$-$	$+$	c	4	16	40	
$+$	$-$	$-$	a	12	24	12	$A + BC$
$-$	$+$	$-$	b	10	8	8	$B + AC$
$+$	$+$	$+$	abc	14	4	-4	$C + AB$

Now we switch signs on everything

A	B	$C = -AB$	Combination	y	1	2	Effect
$+$	$+$	$-$	ab	16	27	41	
$-$	$+$	$+$	bc	11	14	-9	$A - BC$
$+$	$-$	$+$	ac	9	-5	-13	$B - AC$
$-$	$-$	$-$	(1)	5	-4	1	$C - AB$

Obtain all estimates except ABC

Resolution IV Designs

- Can be obtained by folding over resolution III design
- Thus, Resolution IV can be divided into two blocks
- Must contain at least $2k$ runs
- Minimal design if runs equal $2k$
- Determination of defining relation from fold-over III
  - $L + U$ words used as generators
  - L words of like sign
  - U words of unlike sign
  - Combined design will have $L + U - 1$ generators

    All L words

    Even products of U words

Examples

- $I = ABD = ACE = BCF = ABCG$ and $I = -ABD = -ACE = -BCF = ABCG$
- $L = 1$ and $U = 3$
- $I = ABCG = ABD(ACE) = ABD(BCF)$
- $I = ABCE = BCDF = ACDG = ABDH = ABCDJ$ and $I = ABCE = BCDF = ACDG = ABDH = -ABCDJ$
- $L = 4$ and $U = 1$
- $I = ABCE = BCDF = ACDG = ABDH$